National
Institute of
General
Medical
Sciences

BIENNIAL REPORT FY 1988-1989

U.S. DEPARTMENT OF HEALTH
AND HUMAN SERVICES
Public Health Service
National Institutes of Health

R
853
U5461
1988-89

LIBRARY
AUG 1 3 1990
National Institutes of Health

CONTENTS

THE DIRECTOR'S REPORT

This report covers the Fiscal Years 1988 and 1989, during which, while capitalizing on the increasingly exciting scientific opportunities, the National Institute of General Medical Sciences (NIGMS) underwent some changes:

a. In the previous 2-year period, NIGMS played a leading role in developing the initiative regarding the characterization of the human genome. As this initiative evolved, it became clear that its activities are primarily driven by technological developments and by projects targeted to a systematic mapping and sequencing endeavor. Such targeted studies are not in consonance with the mission of NIGMS to support basic, untargeted research. Therefore, having set the human genome initiative on its way, NIGMS agreed with the National Institutes of Health (NIH) that a separate entity should be established. The National Center for Human Genome Research, which was established at the start of FY 1990, will be managed by a staff largely drawn from the skilled program staff of NIGMS.

b. The Institute has continued its leading role in meeting the challenge of providing support for the fundamental biomedical research upon which a new and expanding biotechnology industry depends. Furthermore, the need of this still fledgling industry for research manpower is growing and, because of its expertise and support for research training, NIGMS has been given the responsibility of a new research training initiative in biotechnology.

c. A decision was made by the Director, NIH, to consolidate, into one organizational unit, the major activities of NIH in regard to increasing the numbers of minority researchers. Preparations were made for the transfer of the Minority Biomedical Research Support (MBRS) Program from the Division of Research Resources to NIGMS at the start of Fiscal Year 1990. It is anticipated that this consolidation, along with the provision of supplements to ongoing research grants by funds which are targeted for investigators, graduate students, undergraduates and, in some cases, postdoctoral fellows, all of whom are minorities, will better enable the country to meet the great challenge of the next 10-30 years: the fact that one-third of the entering work force will be minorities. Thus, minority students must be recruited and receive the best training available to enable them to fully participate in biomedical research and to join in assuring continued progress in elucidating the mechanisms underlying human disease.

All four of the Institute's scientific programs have grown in scope. The research objectives outlined by each are becoming more and more complementary and dependent on the interdisciplinary talents of the scientific community. The Institute continues to primarily support investigator-initiated individual research grants and research training endeavors. The new molecular biological techniques, developed primarily through research supported by NIGMS, are being exploited in even greater ways by the more clinically-oriented scientists supported by the categorical NIH institutes.

In the past 2 years, research training has not been fully funded, primarily because stipends and tuition have risen while the funds allocated to such training have remained level. Nevertheless, the NIGMS research training programs remain of the highest quality and serve as the foundation for such training at academic institutions throughout the country. It is gratifying to note that the review procedures used by the excellent staff of the Office of Review Activities have now been recommended for adoption by a new NIH report on research training.

The research contract for GenBank has been most ably administered by Drs. James Cassatt and Irene Eckstrand and is now up-to-date in its data collection. Access to the data has been improved and international coordination is working well. The Human Genetic Mutant Cell Repository started a new 5-year contract period and Dr. Judith Greenberg is continuing to serve as project officer in a most able manner.

As in previous years, FY 1988 and FY 1989 were full of fiscal problems. Despite the uncertainties regarding the availability of funds during the past years and the inability to provide researchers with the levels of support commensurate with the excellence of their scientific endeavors, the superb judgment, skilled oversight, and understanding of the staff have assured that the research supported by NIGMS is of uniformly high quality and that constraints are applied fairly and with sensitivity to the difficulties involved.

It should be emphasized that the attached research highlights are merely representative, and far from inclusive, of the enormous range of superb scientific endeavors supported by NIGMS. It should also be reiterated that <u>fundamental biological research</u> continues to be the single most important factor in rapidly adding to our knowledge regarding the prevention, diagnosis, and treatment of the numerous disorders that continue to plague mankind.

RESEARCH TRAINING

Since its establishment in 1962, NIGMS has had, as one of its principal responsibilities, the support of research training in the biomedical sciences. In fact, through a variety of programs, this Institute provides support for two-thirds of all the predoctoral trainees supported by NIH. In the future, these research trainees are expected to be among the leading scientists who will fill the nation's health research needs and thus carry out the mission of NIGMS and NIH.

All research training support at NIH is authorized by the National Research Service Award Act of 1974. Under this legislation, NIGMS supports predoctoral graduate training through multidisciplinary and multidepartmental research training grants. The intent of these training grant awards is to allow broad and fundamental research training in the basic biomedical sciences. In 1974, NIGMS established four multidisciplinary clusters, namely, cellular, biochemical, and molecular sciences; genetics; pharmacological sciences; and systems and integrative biology. Within these are a wide variety of programs at universities and medical schools throughout the country. Another program, the Medical Scientist Training Program, provides support for students seeking the combined M.D.-Ph.D. degree at outstanding medical schools in the United States. This program is designed to train an exceptional breed of scientist who will play a role in easing the national shortage of clinical investigators.

Postdoctoral support is emphasized somewhat less by NIGMS, and is provided primarily through individual fellowships. This type of award serves several basic science areas and is appropriate when the postdoctoral scientist is working intensely with an established investigator on a particular research problem, with little need for an organized training program. NIGMS does, however, award a small number of postdoctoral research training grants in the more clinically related areas of research training, and emphasizes the selection of M.D. degree holders to be postdoctoral trainees.

The Minority Access to Research Careers (MARC) Program is a unique research training program administered by NIGMS, in collaboration with other NIH institutes. Support is provided for research training in health-related areas for faculty and undergraduate students at minority institutions throughout the country. The most outstanding of these students are subsequently supported by individual fellowship awards to complete their Ph.D. degrees at universities of their choice. A more detailed description of the MARC Program appears in a later section of this report.

In 1987, NIGMS announced a fifth area for multidisciplinary predoctoral training grants--molecular biophysics. The intention of this new program is to provide training which focuses on the application of physics, mathematics, and chemistry to problems of biological structure, primarily at the molecular level. In its initial year, FY 1988, 14 awards were made in this training area. In FY 1989, 13 awards were made. This program is supported by funds allocated by a special appropriation for AIDS-related research and research training.

A sixth multidisciplinary predoctoral research training field, biotechnology, was announced in 1988. These programs emphasize engineering, mathematical and physical methods as they are applied to the production of biological products using the new biotechnology. This training is designed to foster interactions with the burgeoning biotechnology industry. In FY 1989, nine awards were made in this field.

The NIGMS research training budget in FY 1988 was $66.3 million, compared to $64.4 million the previous year. During this period, the number of trainees and fellows (including 100 trainees and fellows in Molecular Biophysics) supported by NIGMS remained almost constant, 3,587 compared to 3,566 in FY 1987. Support for this number of trainees was made possible only by freezing tuition costs at the level paid in the previous year. It should be noted that the actual tuition costs have continued to rise at a rate of about 10 percent.

In 1988, noting that stipend support had not been increased for several years, and had fallen further behind the rate of inflation and the stipends offered by other agencies and foundations, NIH increased stipends for all trainees and fellows, beginning with awards in FY 1989. As a result, the number of trainees and fellows supported by NIGMS (as well as other institutes) was initially reduced, with many research training programs cut significantly below the FY 1988 level. Later in the year, rebudgeting of funds into research training allowed the restoration of most of these trainees and fellows. In summary, the final FY 1989 NIGMS research training budget of $71.9 million supported 3,702 trainees and fellows. During this year, the total number of trainees and fellows supported by NIGMS increased by 115, but that includes an increase of 117 trainees and fellows in the new programs of molecular biophysics and biotechnology.

The total number of trainees and fellows supported by NIGMS has fallen over the last 15 years from over 6,000 to 3,700, with the number of regular predoctoral trainees dropping from over 4,000 to 1,750. As a result, not only has the number of trainees dropped, but the size of each training grant is constrained (predoctoral training grants are currently funded about 20 percent below the level recommended by the National Advisory General Medical Sciences Council). The number of institutions with at least one NIGMS predoctoral training grant has fallen from 102 in FY 1972 to 68 in FY 1988. The increase in average tuition payments has been substantial: $2,700 average per trainee tuition charge in FY 1977 compared to $7,400 in FY 1987 through 1989.

Early in 1989, Dr. James Wyngaardan, then NIH Director, established three task forces to review the NIH biomedical research training programs. In their joint report of October 1989, the task forces suggested several changes in these programs, including:

- a minimum of 2 years of research training to be required of M.D. postdoctoral trainees, with changes in the payback obligation designed to encourage longer training periods;

- improvement of the structure of, and review procedures for, postdoctoral training grants for physicians;

- integration of research training with clinical certification on postdoctoral training grants for physicians;

- periodic examination of new research training needs;

- increased stipends and establishment of a cost of education allowance to replace tuition payments;

- longer career development awards for professional doctorates; and

- increased data collection regarding the training grants being supported and the careers of former trainees and fellows, to permit assessments of program performance.

Finally, the National Research Council, which publishes a report on Personnel Needs and Training for Biomedical and Behavioral Research every 4 years, anticipates that its new report should be available early in 1990. It will contain evaluations, data, and recommendations relevant to the NIH research training programs.

BIOPHYSICS AND PHYSIOLOGICAL SCIENCES PROGRAM

INTRODUCTION

The Biophysics and Physiological Sciences (BPS) Program was established in 1984 by the expansion of the existing Physiology and Biomedical Engineering Program to include aspects of molecular biophysics previously supported by other programs within the Institute. Part of the rationale for this reorganization was the rapid growth of research in the biophysical sciences and the need to provide a focus within the Institute. This rapid growth has been spurred by the application of the techniques of structural biology to the area of computer-aided drug design. Nowhere is this potential usefulness greater than in the elucidation of the structure of the AIDS virus and its encoded proteins for the purpose of designing anti-AIDS drugs. Because of this potential coupled with the national emergency created by the AIDS epidemic, a series of program projects with interdisciplinary teams of scientists was established. This rapid growth has also led to a shortage of scientists trained in molecular biophysics, especially in the area of structural biology. As a result of this shortage a program aimed at training graduate students in the area of biophysics, with an emphasis in structural biology, was established. These latter two areas are discussed in more detail below.

As the title suggests, this Program is heterogeneous. The Biophysics Section focuses on research which results from the application of physical principles to biological systems, as well as the translation of these efforts, through engineering, into practical applications. Instrument development is also supported by this section. The research support in the Physiological Sciences Section is primarily in the area of burn and trauma. This research is directed toward an improved understanding of the total body response to trauma and the molecular mechanisms underlying this response.

In addition to research, the BPS Program supports training through a number of mechanisms. These include predoctoral support through institutional training grants and postdoctoral support through both institutional training grants (for burn and trauma) and individual postdoctoral fellowships. In addition, the NIGMS Medical Scientist Training Program (MSTP) is administered through the BPS Program.

RESEARCH OVERVIEW

BIOPHYSICS SECTION

The dominant theme of this section is determining and understanding the structure, both static and dynamic, and function of biological macromolecules, primarily proteins, nucleic acids, and complexes of the two. In addition, significant progress continues in the development of analytical instrumentation. In terms of information content techniques aimed at determining the structures of biological macromolecules at atomic resolution continue to dominate the field. The most important method is X-ray crystallography; the second is two-dimensional nuclear magnetic resonance spectroscopy (2D-NMR), a field which has emerged in recent years. In addition to the work cited in the highlights found elsewhere in this report, the importance of this technique can be illustrated by the determination of the structure of the *ras* oncogene by Kim and associates and the

protease from the HIV virus determined independently by the laboratories of Navia and Wlodawer. 2-D NMR has proven invaluable for the determination of structures of relatively small macromolecules (molecular weight < 15,000 Daltons) which have proven refractory to crystallization. Although the particular implementation used for structure determination was initiated in Europe, it has received widespread use here. A related technique which bears watching over the next few years is the use of solid-state NMR. This technique is extremely useful for macromolecular samples which can be immobilized, *e.g.* within a membrane. Another technique that offers potential for future development is electron diffraction. This technique requires two-dimensional crystals such as found in oriented membranes and has proved useful in determining the structure of bacteriorhodopsin.

PROTEIN STRUCTURE: Progress in understanding the structure and function of proteins is best characterized by steady progress rather than spectacular discovery. The clear goal remains the ability to predict the folded structure of a protein solely from its primary structure or amino acid sequence. This charge is given special impetus by the initiation of the effort to sequence the human genome. In this project sequences will be determined, but the information stored within these sequences will await the determination of the structures of the coded proteins. To this end, a number of structures have been added to the growing library of three-dimensional structures of proteins determined at atomic resolution, and structures determined by 2D-NMR are now being accepted by the Brookhaven crystallographic database. This area has been greatly facilitated by (1) the ready availability of area detectors that make the collection of data more rapid, (2) the availability of high-energy synchrotron sources that provide a source of X-rays with very high intensity, (3) the mastery of the techniques of structural biology by scientists, for whom structure determination would have been impossible several years ago due to the complexity of the technique, and (4) the use of recombinant DNA techniques that make it possible to produce proteins in large quantities. In spite of the fact that a number of laboratories are beginning to determine structures of proteins by 2D-NMR, it is worth pointing out that the most significant and innovative work is still being done in Europe.

Significant progress is beginning on the determination of the structures of membrane-bound proteins and virus structures. A better understanding of the arrangement of polar amino acids and their arrangement in relation to protein structural motifs is slowly emerging. Site-directed mutagenesis is being used to a greater extent to make well-defined changes in the structures of well-characterized proteins and determine the effects of these changes on the structure, function and stability of these proteins. Work in this particular area has been slow but with the required thoroughness that will provide a large catalog of effects.

NUCLEIC ACID STRUCTURE: The major technical advances of previous years--the ability to produce well-defined fragments of DNA either through cutting at specific sites by restriction enzymes or oligonucleotide synthesis--continue to facilitate progress in the understanding of the structure, both static and dynamic, and function of DNA. One recent group of findings represents in one sense a setback and in another sense an advance in our understanding of the structure of DNA. In 1982 Dickerson solved the crystal structure of the first "long" oligonucleotide, a 12-mer CGCGAATTCGCG, containing B (Watson-Crick) DNA. Based on that structure, many scientists began to

make predictions of the sequence dependence of bond lengths and angles. Within the last year, however, laboratories have crystallized the same oligonucleotide in different crystal forms and found that the structure is highly dependent on the crystal form. This means that the molecule is indeed quite flexible with a number of available conformations of similar energy. Although these studies demonstrate that simple predictions of folded structure of DNA as a function of sequence will not be immediately forthcoming, they point quite clearly to areas of flexibility and rigidity within the molecule.

More significant than the study of DNA itself, however, has been progress in understanding the mechanism of recognition of regulatory proteins for specific sites on their cognate DNA at atomic resolution. In addition to the well-known helix-turn-helix motif associated with many regulatory proteins, the structure of another widely used structural motif, the zinc finger, has recently been elucidated by NMR. Another advance that offers the promise of furthering the understanding of the structure of RNA has been the development of automated synthesis procedures for oligoribonucleotides. These new procedures coupled with new methods for the transcription of RNA from synthetic DNA templates open a whole new field of research on important RNA structures and RNA/DNA hybrids.

INSTRUMENTATION DEVELOPMENT: The development of analytical instrumentation for biological research has continued to evolve, with progress being made in the areas of mass spectrometry, especially with respect to sequencing and the coupling of separation techniques, *e.g.* gas and liquid chromatographies. Significantly, instrumentation developed previously is now being applied to the Human Genome Initiative. One excellent example is the development of pulsed-gel electrophoresis. This technique has been refined to the point where it is now the method of choice for the separation of large pieces of DNA for mapping and sequencing. Another development is the use of mass spectrometry to sequence rapidly segments of proteins. From these segments, probes, which can ultimately be used to sequence a gene of interest, can be synthesized.

AIDS-RELATED SYSTEMS AND THEIR APPLICATION TO TARGETED DRUG DESIGN: The AIDS epidemic has presented a challenge and opportunity to the scientific community. For those in the area of structural biology, especially with its obvious application to the field of computer-aided drug design, the challenge is to take the techniques developed over the last several years and apply them to solve a national emergency--that is to develop antiviral drugs effective against the AIDS virus. To meet this challenge, 2 years ago NIGMS funded interdisciplinary program projects at the following institutions:

University of California, San Francisco
University of California, Los Angeles
Harvard University
Yale University
Smith Kline & French Laboratory
Agouron Pharmaceuticals, Inc.

The idea was to bring together expertise in molecular biology, virology, structural biology,

and medicinal chemistry to meet the challenge described above. In addition to providing direct funding for these efforts, this program has served as an effective catalyst by sponsoring three meetings of scientists working in this general area. A summary of the most recent meeting, held on June 1-2, 1989, is included with the appended highlights. The most important discovery has been the determination of the structure of the HIV protease by two groups independently--Alexander Wlodawer from the Frederick Cancer Center and Manual Navia from Merck, Sharp and Dohme.

PHYSIOLOGICAL SCIENCES SECTION

Research supported in the Physiological Sciences Section of the BPS Program has been directed toward improving the understanding of the complex reactions associated with the physiological and biochemical responses to trauma, including burn injury. In addition, this section supports research projects which focus on the basic biological mechanisms which are involved with behavior and adaptation.

The research programs currently supported in trauma and burn include the study of septic shock, injury-induced alteration of metabolism and the immune response, heat shock proteins, wound healing, and multiple organ failure. The ultimate goal of this research is to save lives, improve treatment, speed recovery and diminish disability associated with traumatic injury. In FY 1989, the Physiological Sciences Section supported 110 research grants, program projects and centers. In addition, NIGMS has planned an international conference entitled "Advances in Understanding Trauma and Burn Injury" for June 21-23, 1990, to be held in Washington, D.C. This conference will represent a major accomplishment for the trauma and burn program and its grantees.

TRAINING OVERVIEW

BIOPHYSICS SECTION

MOLECULAR BIOPHYSICS: In response to a shortage of scientists capable of applying the tools and concepts to the solution of significant biological problems, especially those involving the elucidation of the structure and function of biological systems, NIGMS initiated a predoctoral training program in molecular biophysics in 1988. The BPS program strongly encouraged applications from faculty from departments such as chemistry, physics, and engineering who had an interest in biologically related research, together with faculty in biological science departments whose orientation was to the application of physical methods and concepts to biological systems. Examples of areas of interest included, but were not limited to, studies of biological macromolecules by X-ray diffraction and NMR, theoretical methods applied toward understanding the structure and dynamics of biological molecules, and spectroscopy. The multidisiplinary nature of these programs and the need to ensure that students trained through such mechanisms are well-grounded in tools and concepts of modern biology were stressed.

The first year of the program (FY 1988) a total of 14 awards (which included 103 trainees) were made. The institutions receiving awards were:

Brandeis University
University of Oregon

Cornell University
University of Michigan
Wesleyan University
University of Pennsylvania
University of Washington
University of Illinois
Rice University
University of Minnesota
Columbia University
University of Chicago
Yale University
University of California, San Francisco

The second year an additional 13 awards (54 trainees) were made to the following institutions:

Boston University
Stanford University
Purdue University
University of Texas Southwestern Medical Center
Princeton University
Harvard University
Massachusetts Institute of Technology
Rutgers University
Vanderbilt University

In addition to institutional predoctoral fellowships, the BPS program supports individual postdoctoral fellowships in the area of molecular biophysics. The program especially encourages applicants who are making the transition from disciplines related to molecular biophysics into the area.

PHYSIOLOGICAL SCIENCES SECTION:

BURN AND TRAUMA: Currently there are 13 postdoctoral training grants in the area of trauma and burn research. These grants support 38 trainees, of which 35 hold M.D. degrees and the remaining 3 hold the Ph.D. degree. These numbers represent an increase of three programs and 11 trainees over the last 2 years. The basic research pursued by these trainees includes studies on fibronectin in relation to shock and trauma at SUNY, Albany (Dr. Saba, Program Director); studies on osteogenesis and scar formation at the Massachusetts General Hospital, Boston (Dr. Burke, Program Director); studies on the role of prostenoids in pulmonary response to injury at the University of Washington, Seattle (Dr. Carrico, Program Director); studies on mechanisms of lung injury associated with trauma and sepsis at Brigham and Women's Hospital, Boston (Dr. Hechtman, Program Director); studies on the role of calcium as an insulin secretogogue and the role of nutritional status in the management of trauma at Washington University, St. Louis (Dr. Bessey, Program Director); studies on non-invasive methods for monitoring the healing of fibial fractures at Beth Israel Hospital, Boston (Dr. Silen, Program Director); studies on the permeability of lung capillaries after trauma and sepsis in relation to adult respiratory distress syndrome at the University of

California, Davis (Dr. Blaisdell, Program Director); studies on the biochemical mechanisims and the mediator systems which regulate the inflammatory response to burn injury and trauma at the Scripps Clinic and Research Foundation, La Jolla (Dr. Ulevitch, Program Director); studies on the systemic and local metabolic response to trauma and burn injuries at Rhode Island Hospital, Providence (Dr. Caldwell, Program Director); studies on the immunological and physiological alterations associated with acute injury at the University of South Alabama, Mobile (Dr. Taylor, Program Director); research on septic shock and tissue injury at Oklahoma Medical Research Foundation (Dr. Hinshaw, Program Director); studies on wound coverage and neuroendocrine responses and immunological alterations following burn injury at the University of California, San Diego (Dr. Hansbrough, Program Director); and research in immunology, pulmonary function, nutrition, and wound healing following traumatic injury at the University of California, San Francisco (Dr. Hunt, Program Director).

Individual postdoctoral fellowships are also offered to scientists seeking to improve the understanding of the body's systemic responses to major injury and to foster the more rapid application of this knowledge to the treatment of trauma and burn-injured victims. Of the four trainees who currently hold individual fellowships, two hold the M.D. degree, one holds both the M.D. and Ph.D. degrees, and the fourth holds the D.V.M. degree.

The impact of the training can be measured by the fact that most of the former trainees of the institutional training grants as well as those who have completed individual fellowships are continuing in academic research and teaching or completing residency training.

SYSTEMS AND INTEGRATIVE BIOLOGY: The training program in Systems and Integrative Biology (SIB) plays an important role in the NIGMS training picture. As part of the BPS Program, it has provided predoctoral training in both physiology and biomedical engineering. Programs are multidisciplinary in scope but may emphasize one or the other discipline. The Institute has been insistent that biomedical engineering programs provide a strong biological orientation so that trainees may fully enter the world of biomedical research. There has been similar emphasis on appropriately quantitative training for the more physiologically oriented trainees. The SIB program also has a place in the overall Institute training portfolio. The SIB training concentrates on biomedical problems above the cellular level, in many respects taking up where the cellular, biomedical, and molecular sciences training programs leave off. The most common focus of such integrative training is in the area of neurobiology/neurosciences. Recently, a new discipline has emerged, computational neurosciences, which combines approaches from neurosciences, computer sciences, and engineering.

There has been little change in the portfolio of SIB grants during the past 2 years. As for the rest of the predoctoral programs, budget constraints have resulted in a net reduction in the number of trainees. A major concern of the Institute is the maintenance of high standards in the review of SIB grants.

COMBINED M.D.-PH.D. TRAINING -- MEDICAL SCIENTIST TRAINING PROGRAM: The Medical Scientist Training Program supports highly motivated students with outstanding potential for a research and academic career by providing integrated

scientific and medical training leading to the combined M.D.-Ph.D. degree. The program's goal is to prepare its graduates to function independently in both basic research and clinical investigation.

MSTP grants are made to universities and their medical schools. The 29 institutions that currently receive MSTP support are responsible for the operation of their individual programs, including the selection of students to be supported. Collectively, these 29 programs offer 735 carefully selected trainees a wide choice of graduate programs in the biological, chemical, physical, and social sciences.

Nationally, about 110 of the medical schools in the United States offer opportunities for M.D.-Ph.D. study. There continues to be a noticeable increase in the number of medical schools expressing an interest in participating in the MST program. Several of these schools have applied for MSTP support. Many schools have formalized the organization of their combined-degree programs to provide a central focus/administration for operating a program and for recruiting and advising students. Several have included the participation of all of the relevant graduate training faculty and have changed their graduate school admission practices for combined-degree students so that a student is admitted on a schoolwide basis rather than by individual graduate departments. Some schools have taken the additional step of funding their own programs. As a result of actions such as these, there has been a significant improvement at some institutions in the quality of their combined-degree offerings. The result of these efforts has been that three new, very meritorious MST programs were funded in FY 1987, and a fourth new program was funded in FY 1988.

It comes as no surprise that the cost of the MST program continues to increase. The total cost for FY 1989 was $15.950 million or $21,700 per trainee. All of the participating schools are supporting their individual programs beyond the amounts awarded by NIGMS by supplementing trainee stipends and in many cases underwriting substantial recruitment costs. Several schools have additional combined-degree students who are totally supported by locally derived funds but who are otherwise treated as though they were MSTP trainees.

In the last few years the competition among MST programs to recruit the very best students has increased markedly. An increasing proportion of the undergraduate student population who may be prospective MSTP applicants are becoming aware of the possibility of obtaining a combined-degree education and what the attributes of the individual programs are. In addition, many of the MSTP schools have revamped their programs, taking a more assertive recruitment stance and including program features that enhance the combined-degree training experience. It has become abundantly clear that even the most well-established programs must regularly examine themselves and be willing to revise/modify their mode of operation to stay competitive. Program development is a continuous, neverending process.

APPENDIX I:

SUMMARY: THE THIRD MEETING OF GROUPS STUDYING THE STRUCTURE OF AIDS-RELATED SYSTEMS AND THEIR APPLICATION TO TARGETED DRUG DESIGN

Session on HIV Envelope Glycoprotein and CD4 David Eisenberg

Any complete understanding of the pathogenesis of AIDS will require structural information about the two molecules featured in this session: the HIV envelope glycoprotein, often termed gp160, and the CD4 receptor which permits entry of HIV into lymphocytes. Structural studies of both molecules have been initiated, but detailed atomic results seem much closer for the CD4 receptor. One reason for the better progress on CD4 is the availability of relatively large amounts of pure CD4 receptor. Structural studies of gp160 have been hampered by the lack of experimental material, as well as by the heterogeneity of the small amounts of protein that are available. Consequently, most structural studies of gp160, or its subfragments gp120 and gp41, are on model systems.

Studies of the Envelope Protein

Three reports were given on studies of the transmembrane regions of membrane proteins, with the goal of understanding some of the membrane-related functions of HIV. Engelman *et al.* discussed a model system, consisting of a fusion protein of staph nuclease and glycophorin A. This system is capable of yielding information on the residues involved in dimerization within the membrane. Chemical probes for examining proteins within the bilayer were discussed by Delfino *et al.* These may enable them to identify the region of gp160 that is responsible for virus-host membrane fusion, an important step in pathogenesis.

Another report on models for gp41 was given by Fujii *et al.* They studied the effects on liposomes of synthetic peptides having the sequence of segments of gp41, and found several that are lytic and fusogenic.

The only report on the envelope glycoprotein itself was offered by Schawaller et al. They investigated the subunit composition of gp160, gp120, and gp41 using cross-linking reagents. These glycoproteins appear to be tetramers as components of virus particles.

Studies of CD4

Synthetic peptides from CD4 were studied by Eiden *et al.* They focused on the segment of CD4 consisting of residues 81-92, which blocks HIV infection *in vitro*.

Studies of the carbohydrate structure of recombinant soluble CD4, grown in Chinese hamster ovary cells, was reported by Carr *et al.* These authors used mass spectroscopy to identify the major components of the oligosaccharides.

Two crystallographic studies were reported of CD4-related molecules. The first is of CD4 itself, reported by Liddington *et al.* They have collected data to 5.5 A resolution, including some heavy atom derivative data. This study seems likely to yield atomic-level results before long.

The second crystallographic study, reported by Silverton *et al.*, is of the binding region of the monoclonal antibody OKT4a, which binds strongly to CD4. This antibody blocks the binding of gp120 and inhibits HIV-induced syncytium formation. The Fab was prepared and crystals of it are being sought. Also sought are crystals of the Fab complexed with CD4.

Protease Session Charles Craik

Ten presentations from academic institutions and industrial laboratories revealed some of the excitement and extreme frustration that currently engulfs those working on the proteases, the gag proteins, and the integration proteins of retroviruses. Although the majority of the talks focused on the viral protease, some noteworthy advances were discussed concerning gag and integrase.

Structural proteins that are essential for virion assembly are synthesized as large polyprotein precursors that are subsequently processed by a viral or cellular protease. The matrix, capsid, and nucleocapsid proteins are initially contained within the gag polyprotein precursor. During translation of the HIV genome, a frameshift can occur resulting in a gag/pol fusion which encodes reverse transcriptase and the accompanying RNase H activity and integration protein. The processing of the gag and gag/pol polyproteins is attributed to the virally encoded protease, making the protease an attractive target for chemotherapeutic intervention.

For any serious structure/function analysis of a protein, efficient methods for obtaining adequate quantities of the macromolecule and suitable methods for measuring its activity must be available. Recombinant methods for expressing the protein of interest in bacteria or yeast are undoubtedly the best choices for obtaining sufficient quantities of the viral proteins since: 1) the need for working with live virus is eliminated, 2) *E. coli* and *S. cerevisiae* grow rapidly and are inexpensive hosts, and 3) the heterologous expression system permits the preparation of site-specifically modified proteins that can be designed to test specific postulates. Although clever tricks have been tried including secretion, fusion to another highly expressed or easily purified protein, expression in protease deficient and/or heat shock response deficient strains, and replacement of problematic amino acids with less troublesome amino acids, the protease has remained recalcitrant to overexpression efforts above 1 mg/liter in bacteria or yeast.

Debouck and her colleagues (Smith Kline & French Labs) were one of the initial groups to establish expression of the HIV-1 protease in bacteria. Their system produces the protease as a precursor that self-processes to the mature, active form which migrates as a 99-amino acid polypeptide of approximately 11kD on SDS-PAGE. Large-scale fermentation results in sufficient protein for enzyme purification. The purified protease hydrolyzes synthetic peptides that mimic the junctions between proteins in the gag or gag/pol precursors. The activity of the enzyme is readily inhibited by pepstatin, classifying it as an aspartyl protease. However, all known aspartyl proteases contain two aspartic acids

at the active site and range between 28 and 45 kd. The viral aspartyl protease appears to differ from the mammalian counterparts by dimerizing, with each monomer providing one of the catalytic aspartic acids, as shown by size-exclusion chromatography, ultracentrifugation, and crosslinking studies. Presumably, the low levels of expression reflect the toxicity of the protease. Levine (NIH) attempted secreting the protease from *E. coli* by fusing it to protein A but discovered that intracellular expression was required to avoid degradation. The protein remained soluble and could be purified on a IgG column by virtue of the protein A:IgG association. Appelt (Agouron Phamaceuticals) fused the HIV-1 protease to the highly expressed protein dihydrofolate reductase and recovered the fusion protein as insoluble inclusion bodies in bacteria. Then, by replacing the cysteine residues in HIV-1 with alternate amino acids, C67V and C95M, the variant protease auto-processed and remained soluble. Craik and colleagues (UCSF) resorted to fusing the HIV-1 protease to human superoxide dismutase for intracellular expression in *E. coli* and *S. cerevisiae*. To overcome the requirement for self-processing and to facilitate subsequent purification steps the HIV-1 and HIV-2 proteases were secreted from yeast. By fusing the leader sequence of yeast alpha-factor to the 99-amino acid mature form of the protease, active protease could be readily purified from the extracellular culture media. This system also permits the expression of inactive variants (e.g. D25E, D25N) of the protease since the self-processing event is not required for production of the mature enzyme.

Considering the interests of the meeting attendees, the most exciting presentations dealt with three-dimensional structure determinations of viral proteases. Navia and coworkers (Merck, Sharp and Dohme) used a bacterial expression system that yielded 8 mg. of pure protease from 3500 liters of fermentation media. Crystals that were suitable for X-ray diffraction analysis were grown and heavy atom derivatives were made using mercury and lead compounds. The structure was determined to 3Å resolution by multiple isomorphous replacement (MIR) methods in what appears to have been record time for solution of a structure *de novo*. The three-dimensional structure confirms the identity of the enzyme as an aspartyl protease. The symmetry of the space group suggests that the two monomer components of the HIV-1 protease dimer are related by a crystallographic two-fold axis of symmetry placing the active site aspartic acids within hydrogen bonding distance of one another. The structure is predominantly beta-sheet and an extensive dimer interface is formed by two sets of antiparallel beta-strand interactions.

Wlodawer and coworkers (NCI) elegantly circumvented the complications of working with recombinant HIV-1 protease by using the structure of the Rous sarcoma virus (RSV) protease which they obtained from natural sources in sufficient quantities for crystallographic studies. The structure of the RSV protease was solved at 3A resolution using MIR methods. This provided the essential phases needed for solution of the related HIV-1 protease by molecular replacement. Kent (CIT) then provided 0.6 mgs of pure chemically synthesized HIV-1 protease that was produced in a peptide chemist's tour-de-force. The troublesome cysteines were replaced with α-amino-n-butyric acids to simplify the synthesis. Uranyl and platinum derivatives of the resulting crystals were made, permitting solution of the structure by MIR methods as well as by molecular replacement. The structure determined by Wlodawer differs from that of Navia in the main chain connectivity, the secondary structure, and the topology of the dimer interface regions. The differences between the two structures are currently under intense scrutiny since they are of such importance for the rational design of inhibitors targeted to the active site or the essential dimer interface of aspartyl proteases. The availability of the protease structures

coupled with the various genetic and chemical systems for manipulating the primary structure of the enzyme now allows a detailed analysis of the role specific amino acids play in the substrate recognition and catalysis. This information may then permit the development of an effective antiviral pharmaceutical.

Reverse Transcriptase Session Don Wiley

This year's session demonstrated steady progress toward the crystallization of the HIV-1 reverse transcriptase (RT). The presentations fell into three categories: 1) studies on the mechanism of RT designed in part to determine rate-limiting steps and to assist in the preparation of complexes for future structural study, 2) studies on the physical and biochemical properties of various forms of RT produced from various expression systems aimed at providing fundamental knowledge to guide crystallization attempts, and 3) studies on crystallization and low resolution X-ray diffraction from RT.

Three groups reported crystals of RT. One group reported crystals from the P66/P51 heterodimer and small needle crystals of P51. Another reported small crystals of P66 homodimer and plate crystals of P66/P51 heterodimer. The third group reported crystals of the P66 homodimer, P66/P51 heterodimer, a crystal containing both hetero and homo dimer, crystals with various double-stranded DNA oligomers, and crystals of a series of single point mutations and C-terminally truncated RT's.

To date, the best ordered crystal diffracts to about 6A resolution. X-ray data collection is underway on those crystals, which appear to have many dimers per asymmetric unit.

The number of crystallization conditions screened now number in the thousands; and, most obvious strategies such as crystallization with DNA, crystallization of mutations of RT, and crystallizations with monoclonal Fab's have been attempted.

The general impression was that most of the obvious ideas have been attempted and that while similar ideas may yet yield a better crystal, new approaches should be encouraged.

Tat and Some Research Tools Angela Gronenborn

This last session was devoted primarily to studies on the trans-activator protein tat, including both structural and functional properties of this regulatory protein. Nuclear magnetic resonance (NMR) spectroscopy indicates that the apo-protein (Cd free) is a highly flexible molecule which contains very little ordered structure (L. Mueller). At present, it seems rather unlikely that a structure determination by NMR for this protein can be accomplished, since the high degree of disorder and the presence of multiple conformers prevent a rigorous analysis of the spectra.

Several contributions addressed the question of nucleic acid binding exhibited by the tat protein. Evidence was presented (J. Karn) which implicated purified recombinant tat in

direct sequence-specific binding to TAR RNA. In contrast to this report, experiments using whole cellular systems (R. Gaynor, K.T. Jeang) implicate several cellular proteins in TAR RNA-protein complex formation. It seems clear from all reports that the secondary structure of the TAR RNA plays an important role in the recognition process. Whether tat itself plays any active role or is only a mediator for cellular factors will have to await further clarification. It is most likely that tat has some intrinsic non-specific RNA binding activity which, however, might not be directly linked to its biological function.

Activation and inhibition of trans-activation was studied using synthetic peptides comprising portions of the complete tat sequence (A. Frankel), indicating that the entire protein sequence is needed for the biological function of tat. Inhibition of trans-activation by tat peptides is most likely a non-specific effect on transcription in general and not connected to a tat-linked function.

Evaluating all of the above results it emerges that the HIV trans-activation is still a poorly understood event involving cellular RNA binding proteins which associate with TAR RNA, most likely modulated by an interaction with tat. Additional studies are clearly necessary to further unravel the specifics of this HIV function.

Molecular modeling, primarily focused on mapping the active site of proteins, is being pursued as an initial attempt in inhibitor design (R. DesJarlais). It is anticipated that based on the encouraging experience with aspartyl protease inhibitors, an extension of this approach to the HIV protease will yield useful results.

APPENDIX II

RESEARCH HIGHLIGHTS

Memorandum

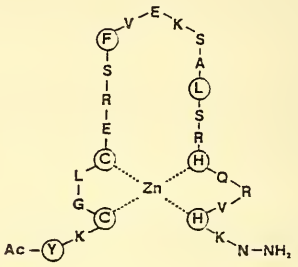
Date November 27, 1989

From Director, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Structural Studies of Zinc Fingers"
R01 GM36643 and R01 GM38794 (Wright, Peter)
Scripps Clinic

To Director, NIGMS

The binding of regulatory proteins plays a major role in the control of gene expression. One of the significant advances in recent years has been the determination of the structures of several of these proteins (in some cases with their bound recognition sequences of DNA) by X-ray crystallography. The structures of the regulatory sequences solved thus far, e.g. *trp*-repressor, the *lambda* headpiece, and *cro* repressor, all demonstrate the by now familiar helix-turn-helix motif. Other motifs, however, have been described recently. These include the leucine zipper motif and the zinc-finger motif. The structure of the latter has, until recently, proved difficult to determine since the segment of the protein that forms a so-called zinc finger is part of a larger structure which has not yet been crystallized. That structure generally contains several putative zinc finger sites. However, recently the structure of a zinc-finger has been solved in the laboratory of NIGMS grantee Peter Wright (GM 36643 and GM 38794) using 2-D NMR.



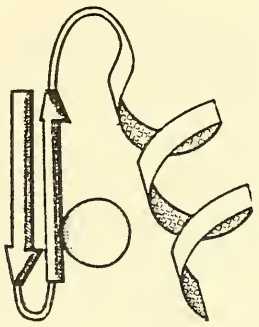
1 Amino acid sequence of finger 31 from *Xfin* drawn to illustrate the finger motif.

2 Structure of the *Xfin31* zinc finger. The overall shape is like a right hand with the thumb representing the amino terminal ß strand.

The zinc finger motif was first identified in transcription factor *IIIA* from *Xenopus* oocytes. The zinc is bound to the peptide either by two cysteine and two histidine residues (*TFIIIA* type) or four cysteine residues (steroid receptors and yeast transcriptional factors). Circular dichroism and NMR have been used to determine that the folding of synthetic single

factors with sequences from *TFIIIa* and *ADR1* are dependent on the binding of zinc.

In Dr. Wright's work a peptide corresponding to the 31st zinc finger from the *Xenopus* protein *Xfin* (denoted *Xfin31*) was synthesized and the structure determined by 2-D NMR. Gel mobility assays demonstrate that the *Xfin31* binds non-specifically in the presence of zinc and does not bind in its absence. A schematic diagram of the proposed complex is shown in figure 1. The three dimensional structure is shown in figure 2. The first ten residues form a hairpin loop that includes the two cysteine residues that bind to the zinc ion. A well defined helix, which includes the two histidine residues that complete the coordination sphere of the zinc ion, extends from glutamine 12 to lysine 24. The zinc ion itself is buried in the interior of the molecule. The overall shape of the zinc finger as shown in figure 2 is like a right hand with the zinc and the carboxyl terminus at the base of the hand and the amino terminus extending like the thumb toward the fingertips.



3 Proposed model for a zinc finger.

Previously proposed models for zinc fingers bear a striking resemblance to the structure determined by Dr. Wright. In this context the best model is that proposed by Dr. Jeremy Berg from Johns Hopkins University (supported by GM 38230). That model was derived from considering sequence homologies to known structures, the critical placement of metal binding groups and the known tendencies of the various amino acid residues to form various secondary structures. Details of that model are shown diagrammatically in figure 3.

The central question not answered by these studies is how the zinc fingers bind to their cognate DNA sites. As noted by Dr. Wright there is evidence that binding occurs in the major groove of DNA and, based on known interactions of other structural motifs with DNA, that binding probably takes place through the helical portion of the zinc finger. The face of that helix contains a high density of positively charged amino acids (lysine 13, arginine 18, arginine 21, and lysine 24) that would facilitate the binding to the negatively charged bases on the DNA and could be involved in specific amino acid/base pair recognition. Since molecules that contain zinc fingers characteristically contain several of these structures, the binding would be enhanced through several contact sites on the protein. This model is consistent with that proposed by Dr. Berg shown in figure 4.

4 Model for the interaction between a protein consisting of tandemly repeated zinc finger domains and DNA based on the predicted structure for the individual domains. The α helix from each domain lies in the major groove of the DNA.

Currently structures of several other zinc fingers are being completed. Similarities and differences among these structures should make a further refinement of the proposed binding to DNA possible, as has been done with the operator proteins. The ultimate definition of the mode of binding to DNA, however, awaits further study.

James Cassatt, Ph.D.

Memorandum

Date December 5, 1988

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Crystallographic Study of Repressor-Operator Interactions, 5 R01 GM 31471-06, Dr. Carl O. Pabo

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS *pcc*
Chief, Biophysics Section, BPS, NIGMS *111*

Sequence-specific DNA-binding proteins play central roles in molecular and cellular biology: They regulate gene expression and serve key structural and catalytic roles in other cellular processes. Crystallographic studies are now helping us understand how these proteins recognize specific sites on double helical DNA and how the bound proteins regulate gene expression and affect DNA conformation.

Recent work by Dr. Pabo's laboratory has led to structure determination, at high resolution, of a repressor-operator complex. His laboratory is studying the repressor from bacteriophage lambda, which regulates gene expression by binding to six different sites in the phage DNA. Each site contains 17 base pairs and is approximately symmetric about the central base pair. The amino-terminal domain of the repressor is the DNA-binding domain, and Dr. Pabo's laboratory crystallized this domain with a 20 base pair DNA fragment that contains the operator site O 1. Using data from these crystals, they have determined the structure of the repressor-operator complex at 2.5 A resolution. This structure gives a detailed picture of the contacts that repressor makes with the DNA. It also is interesting because lambda repressor has been the subject of extensive genetic and biochemical analysis, and there have been many studies of mutations that affect recognition and regulation.

The overall structure of the repressor-operator complex shows that the protein dimer binds by fitting one helix from each subunit into the major groove of the operator DNA (figure 1). The complex is roughly symmetric, and the twofold axis of the protein dimer coincides with the approximate twofold axis of the operator site. The repressor monomer contains five alpha helices. The first four helices form a compact globular domain, with helices 1, 2, and 3 on the face that is closer to the DNA. Helices 2 and 3 form a conserved "helix-turn-helix" unit that has been observed in several DNA-binding proteins. As predicted from Dr. Pabo's earlier structure of the repressor protein, residues in these two helices make many of the contacts with the operator, but the cocrystal structure reveals that residues in the C-terminal part of helix 1 and residues in the loop following helix 3 also make important contacts. As predicted from biochemical studies and modelling, the N-terminal arm wraps around the DNA and makes contacts in the major groove on the back side of the double helix. The operator is fundamentally B-form DNA, although the ends of the operator site bend slightly towards the repressor.

The cocrystal structure reveals unexpected side chain-side chain interactions that are important for recognition and also shows subtle differences between the two halves of the operator site. As predicted from the repressor structure, glutamine 44 (which is the first residue in helix 3) forms two hydrogen bonds with an adenine. However, the cocrystal structure reveals an interesting and unexpected aspect of this interaction: The contact made by glutamine 44 is stabilized by the side chain of glutamine 33 (figure 2). Serine 45 (which is the second residue in helix 3) also makes sequence-specific hydrogen bonds, but the position of the serine in the two half-sites is different. In one half-site the serine hydrogen bonds to the N7 of a single guanine, while in the other site it is shifted and interacts with two neighboring guanines. Asparagine 55, which is in a loop just three residues after the end of helix 3, also make sequence-specific contacts. As predicted, this asparagine hydrogen bonds to a guanine, but lysine 4 makes an interesting and totally unexpected contact with this guanine and with asparagine 55 (figure 2). It had not been possible, from the earlier modelling, to anticipate these side chain-side chain interactions. The changes and corrections emphasize the tremendous structural complexity of the protein-DNA complexes.

The repressor dimer also makes about two dozen hydrogen bonds with the sugar-phosphate backbone. These interactions probably provide a significant portion of the overall binding energy, and the intricate network of hydrogen bonding interactions suggests that they play a key role in positioning the residues that contact the edges of the base pairs. Surprisingly, lysine and arginine are not used to make many of the phosphate contacts. These contacts are made by shorter, polar residues and in several cases by the peptide -NH groups of the protein backbone.

This crystal structure has important implications for other studies of protein-DNA interactions. It clearly shows that several side chains can "cooperate" to recognize a single base (figure 2). Such interactions seriously complicate the use of genetic approaches that try to analyze contacts by swapping segments of different proteins or by deleting sections of a protein. This study also raises interesting questions about the precise role of symmetry in protein-DNA interactions. The different contacts made by serine 45 show that differences in neighboring bases can modulate recognition even when the primary contact site is conserved. This subtle asymmetry also complicates attempts to use biochemical, chemical or genetic studies to test detailed predictions about the interactions between side chains and bases.

Comparing the lambda repressor-operator complex with the known structure of the Eco RI complex (determined by John Rosenberg's group at the University of Pittsburgh) and with the predicted structures of other repressor-operator complexes suggests that there is little prospect of any simple "recognition code". A wide variety of contacts have been observed or predicted, and it does not appear that there will be any simple "rule" telling how a particular

amino acid will be used or how a particular base will be recognized. It appears that recognition, like protein folding, involves cooperative interactions and we need to understand it in fundamental physical terms rather than looking for a simple recognition code.

John C. Norvell

John C. Norvell, Ph.D.

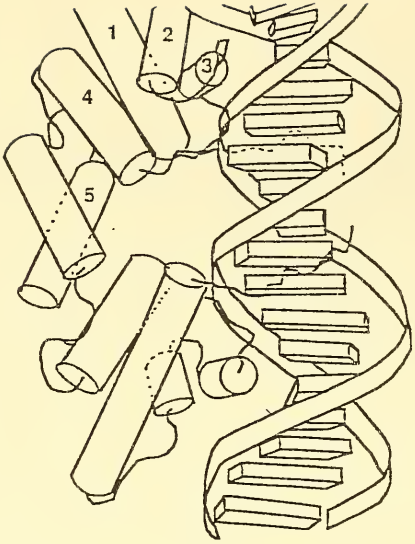


Figure 1

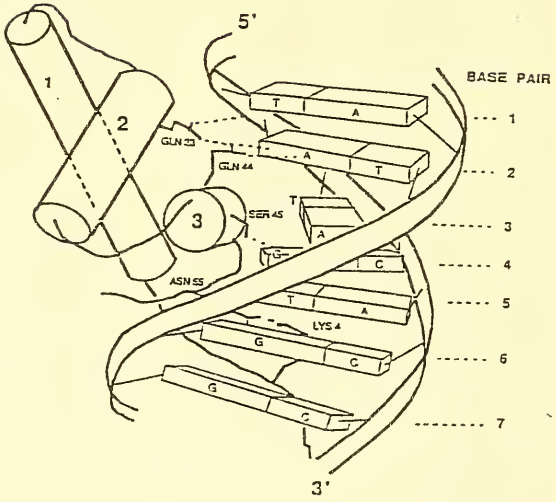


Figure 2

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 24, 1989

From Program Administrators, BPS Program, NIGMS

Subject RESEARCH HIGHLIGHT: Physical and Genetic Studies of Repressor Proteins, 5 ROI GM 22441-13 and 5 ROI GM 35133-03 (Dr. Kathleen Shive Matthews, Rice University)

To Director, NIGMS
Through: Director, BPS Program, NIGMS

Genetic control is an essential function common to all living organisms. In prokaryotes, the ability to alter the expression of genetic information provides flexibility in responding to a continuously changing external environment, and bacteria have been notably useful in defining mechanisms for regulating the transcriptional stage of DNA gene expression. Both positive and negative strategies are employed to regulate initiation of synthesis of messenger RNA, which carries the sequence information of the DNA in a form to be translated into a functional protein structure. Control at this first step in transcribing the genomic DNA to form the mRNA copy yields maximum conservation of cellular energy. Negative control of transcription is exerted by regulatory proteins, termed repressors, which bind to specific regions on the DNA called operators and prevent RNA polymerase from producing mRNA. The ability to respond to specific signals in the cell is provided by the binding of signal molecules to the regulatory proteins. Changes in shape or conformation of the protein accompany binding of these signal molecules, and the consequence is alteration in the affinity of the protein for its cognate operator DNA sequence. Inducible systems are inactive in producing mRNA in the absence of the initial substrate for that pathway due to the interaction of the regulatory protein with its operator binding site. Induction of mRNA transcription for the system occurs in the presence of the substrate due to the interaction of the appropriate signal molecule with the protein and the consequent diminution in binding affinity for operator DNA. In contrast, repressible systems most often apply to biosynthetic pathways, where gene expression is modulated by the end-product. The generation of mRNA for the system is inhibited only when the end-product binds to the regulatory protein and elicits a change which results in increased affinity for the cognate operator.

The lactose and tryptophan repressors for E. coli are negative regulatory proteins which control the expression of the enzymes associated with the degradation of lactose and the biosynthesis of tryptophan, respectively. Dr. Kathleen Matthews and her coworkers have been exploring the mechanism of action of these proteins using a variety of methods. Detailed understanding of genetic control is possible only through insight into the structure and mechanism of action of regulatory proteins. The lactose and tryptophan repressors are available in quantities which allow biophysical and chemical examination. In addition, the coding sequences of the two repressors have been placed in plasmids which allows the manipulation of the amino acid sequence of each protein.

Lactose Repressor Studies: Multiple operator DNA sequences are involved in regulating synthesis of lac mRNA. Two pseudooperator sequences occur at different distances along the DNA from the primary operator. Removal of these operator-like sites results in decreased binding to DNA containing the primary operator alone. This interaction between sites is mediated by the ability of the lac repressor to bind to two operator DNA segments simultaneously, as the protein is a tetramer of identical subunits with two identical operator binding sites. Dual occupancy of these sites provides a mechanism for increasing the affinity for the DNA, but simultaneous binding to both sites is influenced by the topology of the DNA. Negative supercoiling similar to that found in the cell results in significant stabilization of the complex, by a factor of over 1000-fold, if multiple operator sites are present. These results obtained by Dr. Matthews in collaboration with Dr. Robert Wells (GM 30822) of the University of Alabama at Birmingham are reminiscent of effects observed for regulatory sequences which control mammalian gene transcription.

Changes in the amino acid sequence of the lac repressor at specific sites have been made using modern molecular biological techniques. Although the three-dimensional structure of this protein has not been determined due to its resistance to crystallization, aspects of the structure have been derived from homology to sugar binding proteins of known three-dimensional structure (e.g., arabinose binding protein) by Drs. Matthews and F. A. Quiocho (GM 21371). Using this hypothetical sugar binding site (Figure 1) as a basis for selection in the lac repressor, specific amino acids believed to be involved in binding to the sugar signal molecule are being altered. The first of this series of inducer binding site mutants was the conversion of Arg197 to Gly, i.e., the exchange of a large, positively charged side chain for a single hydrogen atom. When purified and characterized, it was found that the Arg197Gly mutant protein binds to operator DNA with an affinity similar to the parent molecule but, as expected based on the model, its ability to bind to the modulator sugar is decreased by 10^4. Sugar modulation of binding to operator DNA is also diminished by about the same factor. Preliminary results characterizing the mutant proteins Arg197Lys and Arg197Leu show that they have effects similar to the glycine mutation, i.e., they bind to the operator with a normal affinity, but exhibit greatly decreased binding to the modulator sugar. This makes sense, since while the operator should be able to accomodate the bulk of these amino acids, they are unable to make the necessary hydrogen bonds to bind to the sugar. These results provide genetic evidence for the participation of Arg197 in the sugar binding process and offer confirmation of the structural model. Thus, homology to proteins of known structure can allow prediction of residues which participate in binding interactions in proteins of unknown three-dimensional structure.

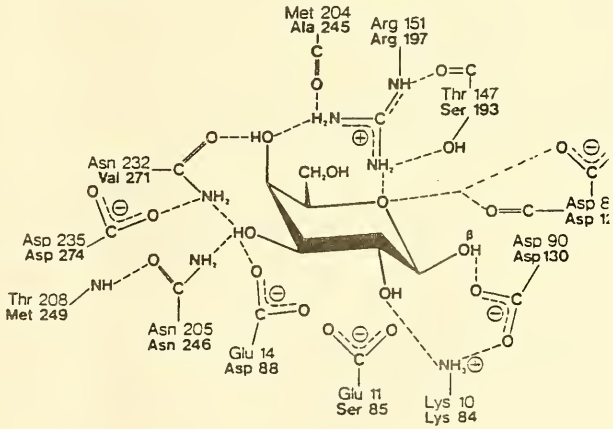
Tryptophan Repressor Studies: Cysteine substitution for serine has been made at several positions within the structure of the dimeric tryptophan repressor molecule. Although this substitution of a sulfur for an oxygen (sulfhydryl for hydroxyl) should elicit minimum perturbation, Dr. Matthews with Dr. Wei-Yuan Chou has found significant effects of this relatively minimal change. At position 67, the greatest effect is on operator affinity, with

minimal alteration of the binding to the end product modulator, tryptophan. Similar substitution at positions 86 and 88 resulted in decreased affinity for both operator DNA and for tryptophan binding. Changes have also been made in residues identified by crystallography (Dr. Paul Sigler, GM 15225) to be in the tryptophan binding site. Some of these amino acid substitutions, e.g., Arg54Leu, Arg84Leu, result in protein with significantly diminished affinity for this modulator and devoid of operator binding since this activity requires the presence of the co-repressor tryptophan. Others, e.g., Thr44Ala, decrease tryptophan affinity moderately and can bind operator in the presence of elevated concentrations of tryptophan.

Effects at the binding site were predictable based on the molecular structure, but those at other sites were not readily anticipated. These experiments illustrate the complexities encountered in generating functional inferences from structure and highlight the delicate character of the active form of proteins, which can be influenced by subtle alterations in interactions.

Studies of these repressor proteins provided detailed mechanistic and thermodynamic information on molecules involved in genetic regulation and form a basis for insight into more complex systems. In addition, the information derived from studies such as these is crucial in arriving at an understanding of the rules of protein folding and the relationships between protein structure and function. Since proteins are the primary expression of genetic material, understanding the relationship between sequence and function is thus of exceptional importance.

Figure 1



Hypothetical sugar binding site in lactose repressor. This site was orginally proposed in 1984 (Sams, C., Vyas, N., Quiocho, F., and Matthews, K., 1984, Nature 310, 429-30) based on homology with the arabinose binding protein (ABP) binding site determined by crystallography (Quiocho, F. and Vyas, N., 1984, Nature 310, 381-6). Residues in bold correspond to lac repressor; residues in normal print correspond to ABP.

Helen R. Sunshine, Ph.D.

John C. Norvell, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 29, 1988

From Director, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Bending and Other Conformational Modifications of DNA (Tullius, Thomas D., GM 40894-04)

To Director, NIGMS

Discoveries of great elegance and importance obviously have great benefits for science. What is not always so obvious is the danger that can accompany the benefits. This is the tendency of the discovery to dominate the imagination so that it is difficult to challenge or modify the theory. Discoveries which can be represented as structures or concrete images particularly present such a hazard. A good example of this kind of behaviour followed the publication of the double-helical Watson-Crick structure of DNA. The image provided is one of a regular double helix with constant rotational and translational helix parameters from one step to the next. This concept of a regular and invariant structure has been implicitly assumed since then. Although evidence of conformational variants of DNA have been obtained, e.g. the A,B,C, and Z forms, each has been described as possessing a regular and invariant behaviour along the chain. In part, this is a result of the kind of experimental data that provided the clues to the double-helical structure. The fibre diffraction data of Wilkins and Franklin gave evidence of the averaged conformational parameters, not the detailed atomic resolution of single crystal X-ray structures. These data did not of course exclude the possibility of local pertubations, but the power of the image generated and its obvious utility largely blocked further thinking in this area. A major breakthrough was the publication in 1981, by Dickerson and Drew, of a single-crystal high-resolution X-ray structure of a double-helical oligodeoxynucleotide dodecamer which showed significant differences in local chain structure and base orientation. In the years since then, a vocabulary of DNA conformation has emerged comparable to that developed for protein structures, and thinking about DNA has included local conformational pertubations. The idea of twist, roll, and tilt, propeller twists, as well as local and long-range bending and changes in the major and minor grooves are terms and concepts now used for a complete conformational description of DNA structure. These are all sequence dependent characteristics and have been shown to be significant in functions such as protein and drug binding, and packing into the nucleosome. A problem, however, is that although the single-crystal X-ray structures have been extraordinarily valuable, they are currently available only for small molecules, slightly larger than one turn of the helix. (2-D NMR has the same restrictions.) Techniques for looking at larger, more "biological" molecules are therefore of great importance. Dr. Thomas Tullius has developed a chemical method of analyzing local DNA structure which can be applied to DNA molecules the size of restriction fragments. This approach promises to add significantly to our knowledge of conformation in DNA.

Several techniques have been used to examine local conformational variability in DNA. In one approach, enzymes have been used which can identify single-stranded regions or hairpin loops. However, the problem with the use of enzymes is that their specificity is not always clear. For example, an S1 nuclease which cleaves a single-stranded region is known to cut only the cytosine-rich strand and not the complementary strand. In this case the specificity is greater than just single-stranded regions, complicating interpretation of the results. Another approach is to use specifically designed molecules such as the chiral cobalt phenanthroline used by Barton to identify Z-DNA regions or the iron (II) complexes designed by Dervan to intercalate into DNA. Tullius' approach is the complete opposite of these - he looked for a completely non-specific cleavage reagent, so that cleavage patterns would be entirely the consequence of local conformation or accessibility, rather than being affected by the identity of the base. He chose the iron (II) EDTA complex, which produces the following reaction:

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + \cdot OH + OH^-$$

Iron (II) was used by both Haseltine and Dervan to cleave following binding of a complex to DNA. Tullius' $[Fe(EDTA)]^{2-}$ complex is negatively charged, so that it will not bind to DNA and only the hydroxyl radical is involved in DNA interactions (Fig. 1).

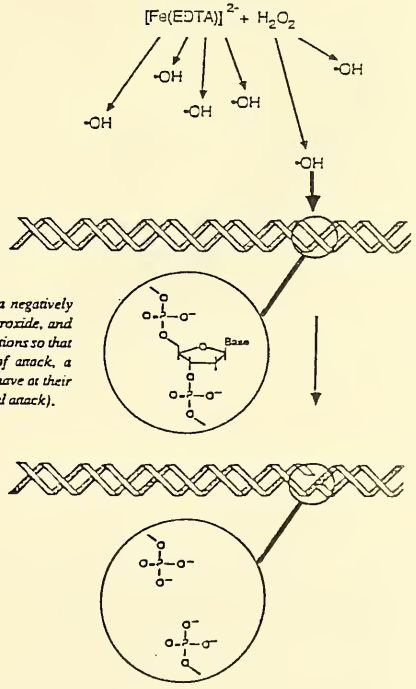


Fig. 1. Hydroxyl radical attacking DNA. The main features of the method are illustrated: a negatively charged complex of iron(II) EDTA sits at some distance from DNA, reacts with hydrogen peroxide, and sends very small projectiles (hydroxyl radicals) toward the DNA backbone. We adjust the conditions so that each DNA molecule is cleaved, at most, once (denoted by the heavy arrow). The site of attack, a deoxyribose, is shown in greater detail. The products of the reaction are DNA molecules that have at their ends the phosphates that were originally linked to the deoxyribose (the victim of hydroxyl radical attack).

Conditions are adjusted so that each DNA molecule is cleaved at most once. The rate of cutting at each site should reflect the kinetic accessibility at that site. The cleavage reaction is assayed by examining the products on a polyacrylamide sizing gel, and densitometry gives a quantitative view of the cutting rates. Since the cleavage is at the sugar, the identity of the base should have no direct influence on the reaction.

The initial studies examined DNA bound to calcium phosphate microcrystals. Rhodes and Klug had observed that the binding of such DNA is in a single rotational orientation. Since only the DNA backbone bonds most exposed to solution could be cut, (Rhodes and Klug used a nuclease as a probe), a periodicity of strong cutting sites was observed which reflected the helical periodicity of the DNA. Tullius noted that this approach would be at least ambiguous for DNA fragments that have defined sequence, e.g. stretches of homopurines or specific palindromes, since enzymes usually have some sequence preference. For such fragments, a cleavage reagent with no specificity, such as iron (II) EDTA, would be necessary. The DNA he studied was a 207bp restriction fragment containing the transcriptional control sequences of the herpes simplex virus-1 thymidine kinase gene. The results (Fig. 2) show a clear periodicity, translating into 10.5 bp/turn, consistent with an average B conformation.

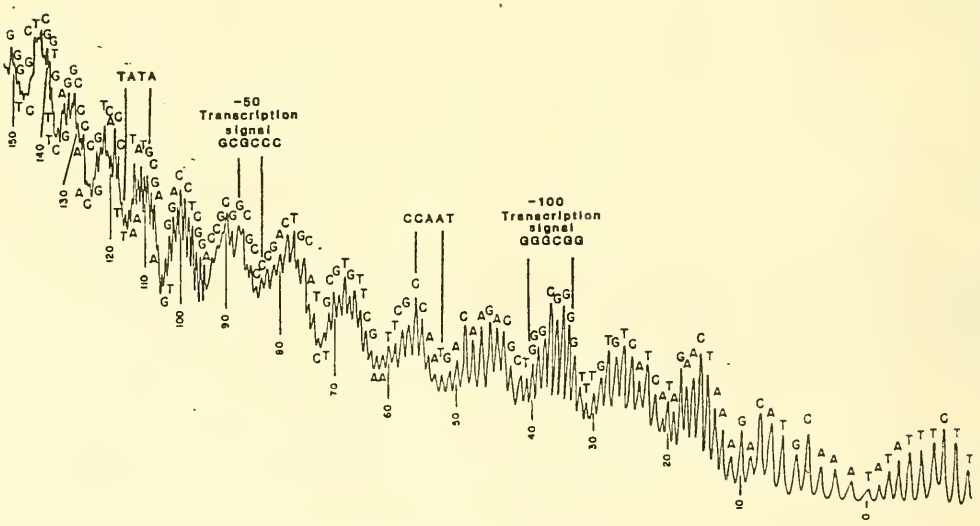


Fig. 2. Helical twist map of the HSV-1 tk DNA fragment. This map was produced by scanning lanes of autoradiographs of electrophoresis gels (such as the one shown in Fig. 1) with a densitometer (0.25-mm slit height, Joyce-Loebl Chromoscan 3 model). The numbers below the map are the number of bases from a position arbitrarily labeled "0." The letter above each sharp peak identifies the corresponding base whose deoxyribose was fragmented by reaction with iron(II) EDTA. The HSV-1 tk map is the result of combining scans from seven separate experiments, which is why the spacing between sharp peaks varies in a nonuniform manner. Scans were spliced together between base positions −1 and 0, 21 and 22, 41 and 42, 73 and 74, 93 and 94, and 115 and 116. Because the splices were made at minima in the pattern, each broad set of peaks is the result of one particular experiment, and the relative intensities of individual sharp peaks within one broad set of peaks can be compared. The intensities of broad sets of peaks cannot be compared with one another, since there was no absolute intensity standard for relating one experiment with another.

This periodicity is a consequence of restricted accessibility to the side of the DNA bound to the crystal surface. The same fragment free in solution shows no periodicity. It is interesting to note that the broad peaks, (reflecting backbone sugars having a lower than expected probability of cleavage by the reagent), are almost all pyrimidine (3'-5') purine sequences. Tullius interprets this as arising from the steric clash of purines on opposite DNA strands, as described by Calladine, resulting in a less accessible DNA surface.

More recently, Tullius has used this technique to examine DNA bending. Bending is a phenomenon that has been associated with biologically important events such as nucleosome positioning and protein and drug binding. It was first observed from the anomolously slow electrophoretic mobility of kinetoplast DNA fragments from trypanosomatids, and confirmed by electron micrographs of this DNA which show 360° bends. How does the DNA bend? These global measurements do not provide such detailed information, although it is known that the bending is associated with poly (dA) tracts. Tullius and his collaborators studied the hydroxyl radical cleavage of a 227 bp fragment of kinetoplast DNA free in solution, and found the highly modulated pattern shown in Figure 3. The sine curve of the cleavage is in phase with the A tracts which themselves occur every 10 to 11 bp, in phase with the helical repeat. The adenine nucleotides at the 3' ends of the A tracts are cut at the lowest frequency, and there is a smooth increase in cleavage frequency to the 5' ends of an A tract. Conversely, the thymine tracts on the complementary strand are cut at greater frequency 5' $\longrightarrow$ 3', with the periodicity of cleavage being offset by 1-2 bases from the adenine position.

The models for DNA bending at poly (dA) tracts have in common the assumption that the effect is uniform at each adenine position. Tullius' detailed analysis shows this is not the case. The key to Tullius' model of bending is the out-of-phase behaviour of the A and T tracts. Note that the reactivity of the adenine in the n^{th} base pair is correlated to the thymine in the (n-1) or (n-2) base pair in the opposite strand. This implicates the minor groove as the determinant of the unusual cutting pattern, since the shortest distance across the minor groove is from one nucleotide to the (n-2)th on the opposite strand. A narrower than usual minor groove would account for the restricted cleavage seen. Unusual narrowing of the minor groove in AT-rich regions has been suggested in other laboratories. Furthermore, Dickerson and his colleagues have shown that netropsin, which binds to the minor groove and forces it open by 0.5-2Å, also bends the DNA back by 8°. The smooth narrowing in the minor groove occuring at poly (dA) tracts, in phase with the helical repeat, could result in a stably bent molecule, and Tullius has shown that even as few as four contiguous adenines can have a structural effect.

This method has great potential in illuminating the relation of sequence and conformation in DNA. Dr Tullius is planning to continue his work on bent DNA, and extend it to the Halliday junctions proposed in genetic recombination and to the tertiary structure of supercoiled DNA in solution.

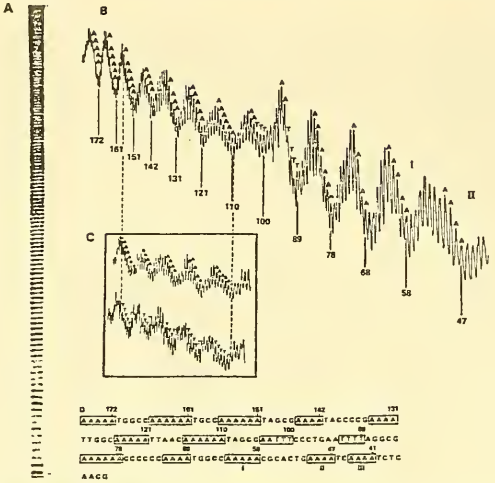


Figure 3. Hydroxyl Radical Cutting Pattern of Kinetoplast DNA

(A) Autoradiograph of hydroxyl radical–induced cleavage pattern of the 227 bp pPK201/CAT SalI–SstI restriction fragment labeled at the SalI site on the 3' end. Dark and light bands, showing the different reactivities of hydroxyl radical toward each individual nucleotide, cause a modulation pattern.

(B) Densitometer scans of autoradiographs of the cutting pattern of the 3'-labeled kinetoplast DNA fragment, with the 5' end of the strand at the left. These scans show a clear sinusoidal pattern, with the adenines in sharp register with the sine wave. Bands are labeled with the base that was attached to the deoxyribose that was attacked by hydroxyl radical. Band assignments were made by reference to a gel lane containing the products of a Maxam-Gilbert guanine-specific sequencing reaction run on the DNA strand. Hydroxyl radical attack of a deoxyribose leaves 5' and 3' phosphate ends, like Maxam-Gilbert chemistry, so hydroxyl radical–induced bands run at the same positions on the gel as Maxam-Gilbert bands. The numbered bases are the least reactive 3' adenines in each tract. A tracts labeled I and II are those referred to in Figure 2. This figure is a composite of scans from two separate experiments that were joined between positions 97 and 98. For this reason the absolute intensities of the groups of peaks cannot be compared with one another. The relative intensities of individual peaks within each group of peaks can be compared, since they result from one particular experiment.

(C) Densitometer scans of both strands of the kinetoplast fragment showing the relation of the cutting patterns of the adenine and thymine tracts. The adenine-rich strand was radioactively labeled at its 3' end, while the thymine-rich strand was labeled at the 5' end. The 5' end of the adenine-rich strand is at the left, while the 5' end of the thymine-rich strand is at the right. The two scans are aligned so that base paired nucleotides lie along vertical lines (such as the broken lines at positions 110 and 156). A tracts labeled I and II are those referred to in Figure 2. The thymines fall consistently on one side of the sine wave, but are offset by 1–2 bases from the adenine positions.

(D) Sequence of the adenine-rich strand of the kinetoplast fragment. Nucleotides are numbered with respect to the SalI end. The same numbering system for the kinetoplast DNA sequence is used throughout the figures and text. Our numbering system is different from that used previously for this sequence (Kitchin et al., 1986). Position 47 in our system corresponds to position 470 in the numbering system of Kitchin et al. Position 172 in our system corresponds to position 345 in the system of Kitchin et al.

Marvin Cassman, Ph.D.

REFERENCES

Tullius, T.D., and Dombroski, B.A., Science 230, 679 (1985)
Tullius, T.D., Trends Biochem. Sci. 12, 287 (1987)
Burkhoff, A.M., and Tullius, T.D., Cell 48, 935 (1987)
Tullius, T.D., and Burkhoff, A.M., in Structure and Expression, Vol 3, DNA Bending and Curvature, Ed. Olson, W.K., Sarma, M.H., Sarma, R.H., and Sunderalingham, M., p. 77 (1988)

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 19, 1988

From Deputy Director, Biophysics and Physiological Sciences Program, NIGMS

Subject "Artificial Restriction Enzyme with User Selected Specificity"
(GM 35724, Dervan, Peter)

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS mac

With the invention of pulsed-gel electrophoresis by Charles Cantor at Columbia University, it has become possible to analyze large fragments of genomic DNA. These analyses are important in the effort to create a physical map of the human genome, a topic of more than passing interest. The ability to analyze large fragments is important because the larger the fragments, the easier it is to overlap them to localize the fragments on a map. With the ability to analyze these large fragments, however, comes a need to produce them. Hence, there is intense interest in searching for restriction enzymes that cut at sites that do not occur frequently in genomic DNA, so called "rare cutters". Dr. Peter Dervan of California Institute of Technology, (supported by GM 35724) has taken a major step forward by chemically synthesizing an "artificial" restriction enzyme, that cleaves double-stranded DNA at a specific site. The important feature is that the specificity of this artificial restriction enzyme can be tailor-made to suit the investigator.

The artificial restriction enzyme synthesized by Dr. Dervan, consists of two parts - one that cleaves the DNA and the other that binds to a specific site on the DNA. The DNA cleaving moity is an iron (II) (ferrous) complex of EDTA (ethylenediaminetetraacetate). Although the reaction is not instantaneous (on the order of seconds) iron (II) compounds are susceptible to autoxidation, that is oxidation to the iron (III) complex by dissolved molecular oxygen. During the autoxidation reaction, oxygen and hydroxyl free radicals are formed as intermediates. These reactive species cleave the DNA.

The DNA binding portion of the artificial restriction enzyme is an oligonucleotide. Previous work had shown that an oligonucleotide covalently coupled with ferrous EDTA bound to and cleaved single-stranded DNA at a specific site whose sequence was complementary to the sequence of the oligonucleotide. The key to the present work was finding conditions under which the oligonucleotide recognized and bound to double-stranded DNA to form a triple helix.

Triple helices have been known since the early work of Felsenfeld et.al. nearly three decades ago. The type of bond that is formed shown in Figure 1. As in the case of the double helix, specificity arises from hydrogen bond formation. In order to determine the optimal conditions for cleavage, studies were carried out under different conditions. The presence of added cations, such as millimolar concentrations of $Co(NH_3)_6^{3+}$ or spermine, is essential to shield the negative charges of double helical DNA from the negative charge of the approaching oligonucleotide. The addition of organic solvents also enhanced cleavage, presumably because when a triple helix is formed there is a slight conformational change of the double helix from B-DNA (Watson-Crick DNA) to A-DNA, a more tightly wound form of DNA that is less highly hydrated.

Although many experiments were done to demonstrate the efficiency and specificity of the cleavage, the key experiment was to cleave a plasmid PDMAG10. This plasmid was converted from circular DNA to linear DNA by cleavage with the restriction enzyme StyI. The linearized plasmid was then cleaved with an artificial restriction enzyme synthesized by linking ferrous EDTA to the oligonucleotide shown in Figure 2. Also shown is the proposed binding of the oligonucleotide and the gel patterns that show conclusively that this compound cleaves the plasmid efficiently at one and only one site.

In summary, although much work remains to refine details, Dr. Dervan's work represents a major scientific advance that will have a significant impact to initially physically map and ultimately sequence the human genome.

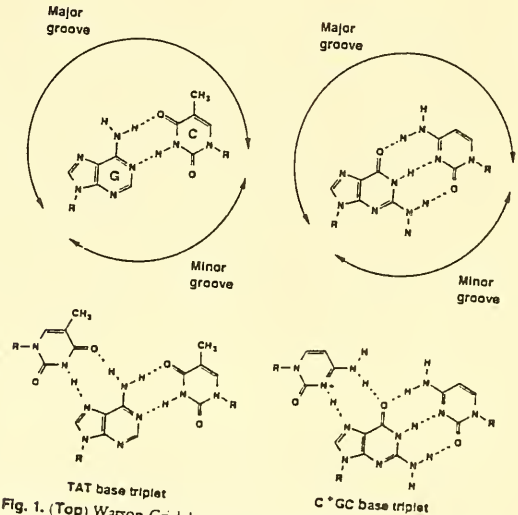


Fig. 1. (Top) Watson-Crick base pairs. (Bottom) Isomorphous base triplets of TAT and C+GC. The additional pyrimidine strand is bound by Hoogsteen hydrogen bonds in the major groove to the complementary purine strand in the Watson-Crick duplex.

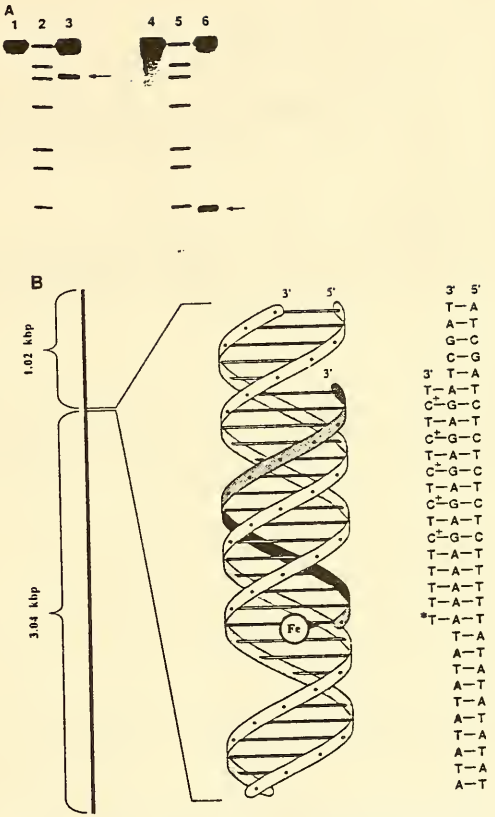


Fig. 2. (**A**) Double strand cleavage of plasmid DNA analyzed on a nondenaturing 0.9 percent agarose gel. (Lanes 1 to 3) Plasmid pDMAG10 linearized with Sty I and labeled at the downstream end of the restriction site with [α-^{32}P]ATP. (Lanes 4 to 6) Same plasmid with the other end-labeled with [α-^{32}P]TTP. Cleavage conditions: ^{32}P-labeled DNA plasmid, 100 m*M* NaCl, 1 m*M* spermine, 25 m*M* tris-acetate, *p*H 7.0, 100 μ*M* (bp) sonicated, deproteinized calf thymus DNA, 5 μ*M* DNA-EDTA·Fe(II) **9**, 25 μ*M* Fe(II), and 2 mM DTT. Cleavage reactions were run for 16 hours at 0°C. (Lanes 1 and 4) Controls containing no DNA-EDTA·Fe(II) **9**. (Lanes 2 and 5) DNA size markers obtained by digestion of Sty I linearized pDMAG10 with Eco RI, Pvu I, Sal I (both ends labeled), and Xmn I labeled with [α-^{32}P]TTP: 4058 (undigested DNA), 3338, 2994, 2368, 1690, 1460, 1064, and 720 bp. (Lanes 3 and 6) DNA-EDTA·Fe(II) probe **9** at 5 μ*M* added. Arrows indicate 3.04-kb (lane 3) and 1.02-kb (lane 6) cleavage fragments. (**B**) (Left) The course resolution cleavage pattern from (**A**). (Center) Simplified model of triple helix complex between the Hoogsteen-bound DNA-EDTA·Fe(II) **9** at single site within 4.06 kb of plasmid DNA.

James C. Cassatt, Ph.D.

Memorandum

Date July 15, 1988

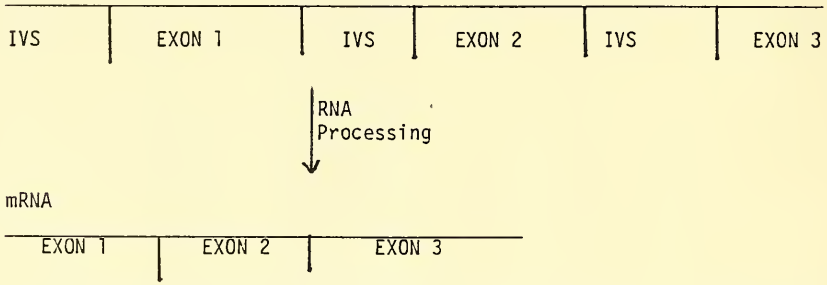
From Program Administrator, BPS, NIGMS

Subject Structure of a pre-mRNA branch point/3'splice site region
5 R37 GM 20168-15 (Redfield) and 5 R01 GM 35490-04 (Green)

To Director, NIGMS
Through Director, BPS Program, NIGMS ZC

One of the fundamental processes in molecular biology is the processing of the primary transcript, that results from the transcription of DNA, to form the intact messenger RNA which is ultimately translated into protein. This process is shown schematically below

Primary Transcript



In this processing step the intervening sequences (IVS) or introns are excised and the exons joined to form the mRNA which is translated into protein.

The study of mRNA processing has been the major focus of several laboratories, and elegant mechanisms have been suggested. Regardless of the mechanisms, however, the structure of the RNA at the splicing junction is presumed to play a significant role. Recently, Merit Awardee Alfred Redfield of Brandeis University (supported by GM 20168) and Michael Green of Harvard University (supported by GM 35490) have proposed a structure for the 3' splice site between the first intron and second exon of human beta-globin (IVS1). Diagramatically this sequence is shown below:

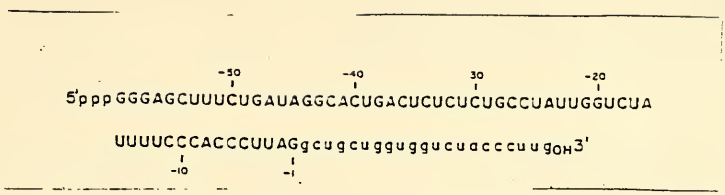


Figure 1: The 3' splice site between the first IVS and second exon of human beta globin. Nucleotides designated by upper case letters are part of the IVs and by lower case are part of the exon.

The molecule was made by transcribing the appropriate portion of the beta-globin gene using the phage T7 or SP6 polymerase systems. Nuclease mapping was done by digesting the above transcript with S1 nuclease, which cleaves at single stranded regions of RNA and cobra venom V1 which cleaves at double stranded regions. The use of these two nucleases enabled Green and Redfield to map the single and double stranded regions respectively.

Nuclear Magnetic Resonance (NMR) is well suited to investigate structures of small RNA molecules as has been demonstrated with tRNA. Specifically the peaks associated with the protons involved in base pairing can be identified because they disappear when the secondary structure is melted on heating. These peaks can be associated with particular types of bases though the use of the Nuclear Overhauser Effect (NOE). In the above molecule it was possible to show that the sequence A.U/G.C/G.C is formed several times within the molecule by the identification of the base-paired hydrogens. Moreover, when the molecule is heated the peaks assigned to hydrogens associated with base pairing disappear in two distinct groups, indicating at least two distinct regions in the molecule.

Combining the results from the nuclease experiments with those from the NMR experiments leads to the following model:

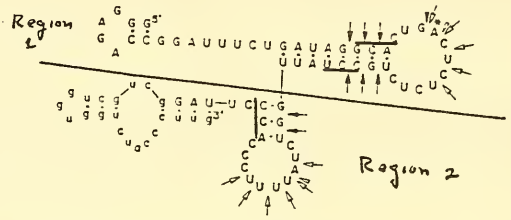


FIG. 2. Structural model of the branch point/3' splice site portion of the human β-globin IVS1. Intron sequences are in uppercase letters; exon 2 sequences in lowercase letters. Sites of primary nuclease cleavages are marked: V1 by closed, tailed arrows; S1 by open, tailed arrows;

The model is consistant with the differential cleavage postions exhibited by the two nucleases. In addition it has two distinct regions as indicated and three A.U/G.C/G.C base pairs shown by the solid lines.

As the authors note "these studies show that in solution the RNA adopts a distinct structure. However, it is not clear if this structure is maintained in the presence of the associated proteins and snRNPs of the splicing complex or what its function would be... [However] the RNA structure could function as a recognition element or could influence the kinetics or efficiency of splicing." In future work the authors hope to vary elements contained within the transcript and examine the effect on both the structure and function to determine the relationship between the two.

James C. Cassatt, Ph.D.

Memorandum

Date August 7, 1989

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Crystallographic Studies of Wild-type and Mutant forms of Staphylococcal Nuclease, GM 36358-03, Dr. Eaton E. Lattman, Johns Hopkins Univ.

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS
Chief, Biophysics Section, BPS, NIGMS

Staphylococcal nuclease (nuclease) is a protein with a long history as a model system in studies of protein structure and function. Now that the technique of site-directed mutagenesis has given rise to an explosion of different mutant forms of many proteins, nuclease is once again proving useful as a molecular laboratory for the study of the effcts of amino acid substitutions on the stability and function of proteins.

Dr. Lattman's laboratory is engaged in a systematic study of the effects of mutations upon the three-dimensional structure of staphylococcal nuclease. Recently, the structure of wild-type nuclease, complexed with calcium ion and the potent inhibitor 3',5'-deoxythymidine bisphosphate (pdTp), has been refined at high resolution in the laboratory. This newly refined structure provides both new insight into the enzyme's mechanism and a very accurate benchmark structure against which structures of mutant forms of the protein may be compared.

In the active site of the wild-type structure, the details of protein side chain configurations, ligand binding, and solvent architecture are clearly shown (Figure 1). The inhibitor molecule is seen to bind in a cleft, with the aromatic ring of the nucleotide shielded from solvent. Interestingly, no specific interactions occur between the nucleotide base and the protein; rather, the hydrogen bonding groups on the base are linked to the protein by bridging water molecules. This provides an explanation for the lack of base specificity which is associated with the hydrolysis of nucleic acids by staphylococcal nuclease. Presumably, both types of pyrimidine bases can be accommodated by rearrangements of the bridging waters, while purines can be fit in by removing some of the water molecules which "pad" the binding cleft. The 3'- and 5'-phosphates are observed to be part of a tightly-interconnected web of hydrogen bonds and ionic interactions involving protein side chains, the metal ion, and solvent molecules. The active site structure is compatible with the mechanism proposed a decade ago by Cotton, Hazen, and coworkers. Glutamate-43 was suggested to act as a general base, abstracting a proton from a water molecule that can then attack the scissile bond, which in this structure is modeled by the phosphoester bond of the 5'-phosphate of the inhibitor. Two water molecules which are candidates for the nucleophile, or specific attacking group for the reaction, are clearly seen in the electron density map. On the basis of the refined structure, the identity of the general acid can clearly be assigned as arginine-87.

Dr. Lattman hopes to clarify one question that emerges from examination of the refined wild-type structure: two lysine side chains from an adjacent molecule in the crystal lattice protrude into the enzyme's active site, forming strong specific contacts with the phosphates of the inhibitor molecule and entering into a salt bridge with glutamate-43, the putative general base. The mechanistic conclusions drawn from the structure are based on the assumption that the structure of the ternary complex of enzyme, metal, and pdTp is a good model for the structure of the productive complex of enzyme, metal, and substrate. If the interaction of the symmetry-related lysines with the pdTp distorts the inhibitor's binding in the active site, then this assumption may not be valid. The best way to test this point is to determine the structure of a complex of metal and pdTp with a mutant form of the protein in which the offending lysines have been altered, perhaps to alanines or glycines. Dr. Lattman's laboratory is preparing to attempt just this.

Applications of site-directed mutagenesis to protein structure-function studies has led to literally thousands of mutant forms of hundreds of proteins. Structural information on mutants has been accumulating, albeit at a slower rate. In the structures that have been seen to date, the altered side chain(s) are incorporated into the folded proteins with little or no structual perturbations beyond the immediate vicinity of the mutation. However, a recent structure of a mutant of staphylococcal nuclease determined by Dr. Lattman demonstrates that this is not a general rule, and that seemingly conservative mutations can cause serious unexpected changes in the structure of a protein. The structure of the glutamate-43→ aspartate mutant form of staphylococcal nuclease (nuclease E43D) has been determined and refined at high resolution. A number of interesting points emerge: 1) The most likely candidate for the nucleophilic water is missing from the mutant structure, indicating that the role of glutamate-43 involves positioning the nucleophile, as well as acting as a general base. 2) A flexible surface loop which adjoins the active site cleft is seen to adopt a drastically different conformation in the mutant structure than in the wild type; the cause of this change in the loop conformation has been traced to the interaction of the mutant side chain with the protein backbone in the loop. 3) Three side chains which are packed into the hydrophobic core of the molecule, some 15-20 A away from the site of the mutation, are seen to adopt different conformations in the mutant and wild-type structures. Since the tight packing of the hydrophobic internal regions of protein is thought to make important contributions to the stability of the folded protein, these changes are potentially of great interest. 4) Comparison of the crystal structure of the mutant protein with solution NMR data obtained in the laboratories of Dr. Gerlt at the University of Maryland and Dr. Bolton at Wesleyan University indicates that the conformation of the protein in solution differs from that of the protein in the crystal lattice. Work is underway to elucidate the precise nature of the differences which can be observed.

Thus, the elucidation of the structue of nuclease E43D has provided unexpected rewards. This structure counsels caution to those who would introduce mutations into molecules and then attempt to interpret their effects on the basis of the wild-type structures, assuming that nothing untoward has occurred except near the mutation. It also suggests that, in addition to causing unexpected rearrangements of the local structure in the vicinity of the mutation, altered side chains can sometimes cause changes at sites quite far removed from the site of mutation. The current state of the theory of protein stability precludes the prediction of the sorts of changes seen in this mutant. The elucidation of more mutant structures may enable experimenters in Dr. Lattman's laboratory to develop an empirical model of how changes in amino acid sequence affect the folding and stability of the staphylococcal nuclease molecules.

John C. Norvell

John C. Norvell, Ph.D.

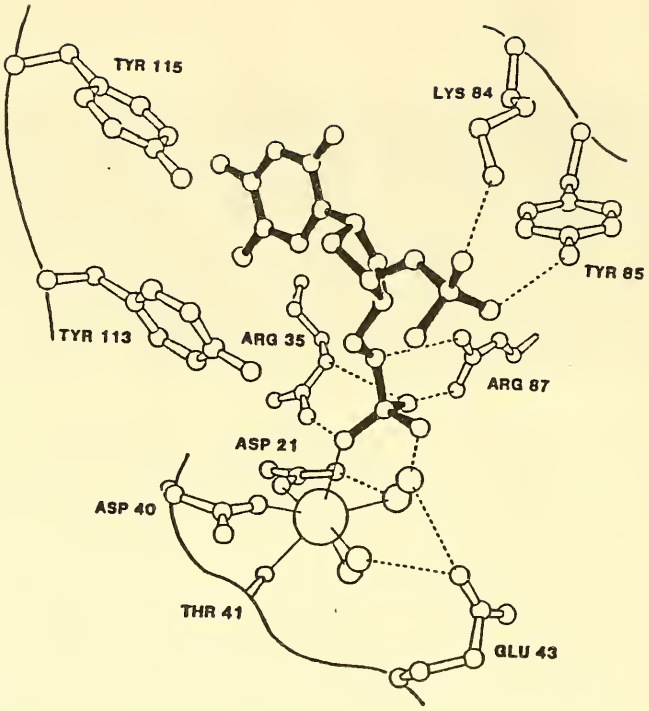


Figure 1 Schematic drawing of the active site of wild-type staphylococcal nuclease. Protein side chains are shown by light bonds, while the pdTp molecule is in dark. The calcium ion is shown as the large sphere below the inhibitor molecule. Also shown are the three inner sphere water ligands of the calcium ion and the water molecule bridging Glu-43 and the 5'-phosphate of the inhibitor (this bridging water is the putatitive nucleophile in the hydrolysis of phosphoesters).

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 14, 1989

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Structure, Function, and Stability of T4 Lysozyme, 5 ROl GM 21967, Dr. Brian W. Matthews, University of Oregon

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS
Chief, Biophysics Section, BPS, NIGMS

Dr. Brian Matthews and his associates at the University of Oregon are using the lysozyme from bacteriophage T4 to test ways in which the stabilities of proteins might be improved by genetic engineering. Such studies are also making it possible to define the contributions that different types of interaction (H-bonds, hydrophobic interactions, salt bridges, etc.) make to the stability of proteins. It is expected that these studies will be useful in stabilizing biological molecules that are used in medicine (e.g. vaccines), and in designing improved enzymes for use in the pharmaceutical and food industry.

There are at least in principle, many ways in which the stability of a protein might be enhanced. Examples include (a) the use of hydrophobic interactions, (b) improvements in hydrogen bonding, (c) the introduction of salt bridges, (d) the introduction of substitutions that decrease the entropy of unfolding, (e) the removal of strain, (f) the stabilization of α-helix dipoles, and (g) the introduction of disulfide bridges. The last two of these methods have been found to be most successful in increasing the stability of T4 lysozyme (as measured by higher melting temperature): the introduction of charged groups to alter helix-dipole interactions and the introduction of non-native disulfide bridges.

(1) Helix-dipole Interactions

In an α-helix the amide groups point toward the amino end of the helix whereas the carbonyl groups point toward the carboxyl terminus. This results in an effective positive charge at the N-terminus of the α-helix and a negative charge at the C-terminus. Therefore, if a negatively charged group could be introduced close to the N-terminus of an α-helix it should interact favorably with the positive charge and stabilize the helix. Similarly, the introduction of positive charges close to the C-terminal ends of α-helices should also improve stability. Recent experiments from Dr. R. Baldwin's lab at Stanford and Dr. Matthews' group show that this is indeed the case.

In the case of T4 lysozyme, initial experiments have focused on the introduction of aspartic acids at or near the amino termini of α-helices. Two such substitutions, Ser 38 $\rightarrow$ Asp and Asn 144 $\rightarrow$ Asp, were both found to increase the melting temperature of protein by about 2°C at pH values where the introduced aspartates were negatively charged. The double mutant was found to increase the melting temperature by about 4°C. A related substitution, Asn 144 $\rightarrow$ Glu has also been constructed and found to yield

essentially the same increase in stability as the replacement with aspartic acid. This therefore seems to be a rather general way to increase protein stability.

Structural studies of the wild-type and mutant lysozymes indicate that the stabilization is due to generalized electrostatic interaction of the introduced aspartic acid side chain with the positive charge at the end of the α-helix, and does not require precise hydrogen bonding to the terminal amino groups. In the case of the Asn 144 $\rightarrow$ Asp substitution, for example, neither the Asn or Asp side chain makes any hydrogen bonds to the end of the helix. Because precise hydrogen bonding is not required it greatly simplifies the design of stabilizing substitutions.

(2) Disulfide Bridges

A number of attempts in several different laboratories have been made to increase the stability of proteins by the introduction of non-native disulfide bridges. Surveys of disulfide bridges in known protein structures show that the geometry of such bridges is very restricted. For this reason it is often difficult to find suitable pairs of residues in a protein that can be linked by a disulfide bridge without concomitant introduction of strain. Although polymer theory indicates that genetically engineered disulfide bridges can increase protein stability, experience to date shows that in many cases the engineered protein is not more stable than wild type.

In an attempt to elucidate general principles relevant to the design of disulfide linkages, five different bridges have been introduced into phage lysozyme (Figure 1). (One of these bridges has already been reported by R. Wetzel at Genentech.) Three of these bridges linking, respectively, residues 3-97, 9-164 and 21-142, increase the melting temperature of the protein by 5°,7° and 11°C at pH 2; the other two bridges destabilize the protein relative to wild-type lysozyme. In each case the oxidized (crosslinked) form of the protein is more stable than the reduced (noncrosslinked) form. Lysozymes in which two or three of the stabilizing disulfide bonds are incorporated into the same molecule are found to be "super stable", with melting temperatures up to 23°C higher than that of the wild-type protein.

The three disulfide bridges that are most effective in stabilizing T4 lysozyme have the largest loop sizes (ie, length of the amino acid chain between the ends of the disulfide bridge). Consistent with theoretical expectation, this suggests that a large loop size is a desirable attribute of engineered disulfide bridges that are intended to maximize protein stability.

When possible, disulfide bridges should also be introduced at sites at which the conformation of the native protein is geometrically compatible with the known requirements for formation of an unstrained disulfide bridge. Bearing in mind, however, that such sites are rare, it may be necessary to choose a

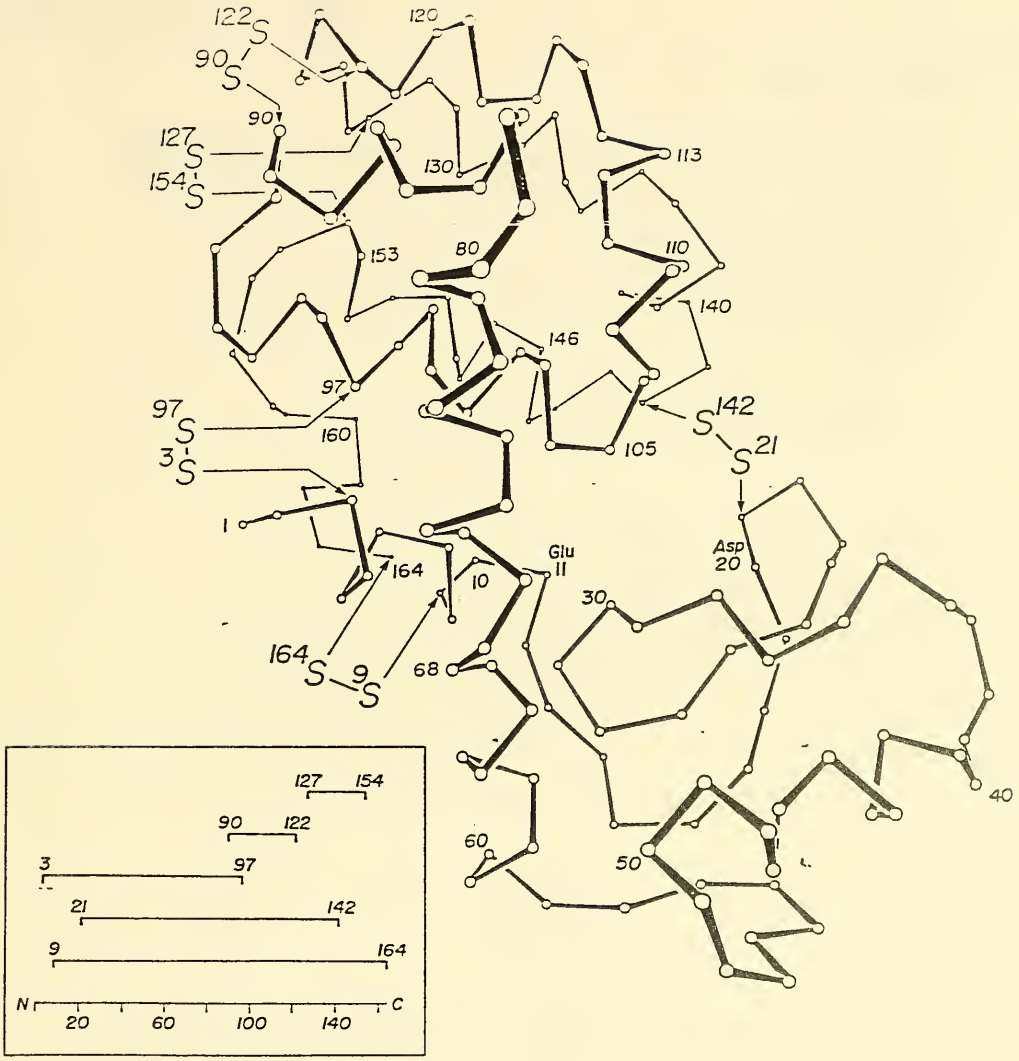


Figure 1. Locations of five disulfide bridges that have been engineered into T4 lysozyme. The lengths of the loops formed by these bridges are shown schematically in the insert.

site that is less than ideal. In the case of T4 lysozyme, the two bridges that are most effective in stabilizing the protein (9-164 and 21-142) are introduced into a flexible part of the structure. The use of such flexible sites could be another general attribute that is desirable in designing disulfide bridges in general.

Thus, substantial progress is being made in developing general ways to increase the stabilities of proteins. One of the encouraging developments is that proteins can be stabilized but still retain their full biological activity.

John C. Norvell

John C. Norvell, Ph.D.

Memorandum

Date August 18, 1988

From Acting Chief, Biophysics Section, BPS, NIGMS

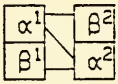
Subject Research Highlight: Ligand Binding to Hemoglobin: A Three (or more) State Model (Ackers, G. K., GM 24486; Johnson, M.L., GM 28928)

To Director, NIGMS
Through: Director, BPS Program

Hemoglobin is a tetrameric protein that consists of two α and two β subunits, each of which contains a heme group (Fe protoporphyrin IX). The heme group can bind to a molecule of oxygen or other ligand, such as CO, NO, CN^-; the ligated form of the heme can also be mimicked by substitution of the heme iron for a different metal.

The binding of oxygen and other ligands to hemoglobin has long been explained by a model in which the hemoglobin tetramer acts as a two-state molecular switch for control of ligand affinity. This concept arose from the allosteric model of Monod, Wyman, and Changeux which specifies two conformational forms of the molecule, each with a different thermodynamic constant for ligands which bind to the heme. It was greatly reinforced by crystallographic studies of Perutz which showed two distinct structural forms for unligated and fully ligated states of the tetrameric hemoglobin molecule. The deoxy crystallographic form is considered identical to the low affinity, or T-state form of the Monod-Wyman-Changeux (MWC) model, and the fully oxygenated crystallographic form is the high affinity, or R-state form of the molecule.

Topographically, the four subunits are arranged as follows:



There are two types of intersubunit contacts within each tetramer, the $\alpha^1\beta^1$ (or $\alpha^2\beta^2$) interfaces and the "$\alpha^1\beta^2$ contact region," which contains three pairwise contacts, $\alpha^1\beta^2$, $\alpha^2\beta^1$, and $\alpha^1\alpha^2$. The two β chains are not in contact. When oxygen or other ligands are bound at all four sites, major structural changes occur at the $\alpha^1\beta^2$ intersubunit contact region, while the $\alpha^1\beta^1$ and $\alpha^2\beta^2$ contacts do not undergo detectable alterations.

When a ligand such as O_2, CO, CN^- or NO binds to the heme, a total of ten ligation states are possible. These are determined by the number of ligands bound and their distribution among the four subunits, taking into account statistical degeneracies due to symmetry. The 10 ligation states are shown in Figure 1.

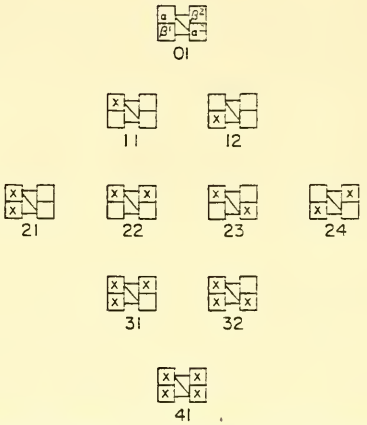


Figure 1 Topographic representation of the ten ligation states of tetrameric hemoglobin. Index numbers provide species designation.

Ligand binding to hemoglobin is highly cooperative, and the population of molecular species is dominated by the two end states, even at half saturation with oxygen (Figure 2).

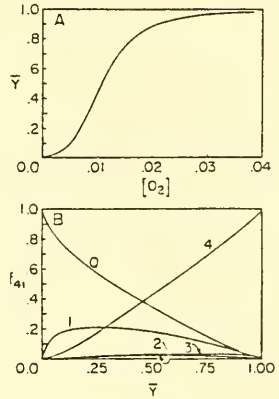


Figure 2 Hemoglobin oxygenation isotherm (*A*), and population map (*B*) for species with different degrees of ligation. [O_2]: Concentration of dissolved oxygen in arbitrary units.

The experimental difficulty of studying the functional and structural properties of intermediate states of ligation, the weight of the structural evidence, and the fact that most of the known phenomena regarding ligand binding to hemoglobin could be explained by this model, have led to the virtual absolute acceptance of the two-state model.

In 1985, however, Dr. Gary Ackers and his colleagues at John Hopkins University (GM 24486) were able, for the first time, to experimentally resolve the Gibbs free energies for all eight intermediate ligation states of hemoglobin. Analysis of these data has led to a reassessment of the two-state molecular switch model for ligand binding to hemoglobin.

The experiments to measure the Gibbs free energy of all ten ligation states were carried out using CN-met hemes as the ligated form of the various subunits and were subsequently repeated, with similar results, for tetramers in which individual heme irons were replaced by Mn(III) as a functional analog for the ligated form, and in molecules where Mn(II) was a model for the unligated subunits while the remaining Fe(II) subunits were ligated to CO. It is necessary to use these analogs of O_2 binding due to the highly labile nature of the O_2 ligand.

In a molecule with more than one binding site for a ligand, site-site interaction may contribute to the free energy of binding. The additional free energy reflecting such site-site interactions is defined as the "cooperative free energy" for the process; cooperativity (i.e., heme-heme interaction), a thermodynamic property, is simply the deviation in free energy of ligand binding from that which would obtain for the same sites if they were ligated independently. The low abundance of intermediate states shown in Figure 2 for O_2 binding to hemoglobin is not an indication of their unimportance to the mechanisms of molecular switching, but merely a reflection of the highly cooperative nature of the system. The free energy, ΔG_i, for reacting i moles of a ligand, X, with the hemoglobin molecule is given as

$$\Delta G_i = i\Delta G_X + \Delta G_C,$$

where $i\Delta G_X$ is the intrinsic binding free energy for the same reaction in the absence of site-site interactions, and ΔG_C is the cooperative free energy, or additional free energy that reflects the site-site interaction.

The central aim of Dr. Ackers work is to: 1) determine the number and distribution of cooperativity states among the ten ligation states of Figure 1 (each distinct value of ΔG_C in the entire set of ligation states defines a separate cooperativity state); and 2) to relate these to the molecular mechanism of hemoglobin ligation (and particularly to the postulated two-state mechanism).

Dissociated dimers ($\alpha\beta$) bind ligands noncooperatively (i.e., with the same affinity as an isolated pair of α and β subunits). Since energetic differences between molecular species are independent of pathway, it was possible, and more convenient, to obtain the cooperative free energy, ΔG_C, for any ligation state by determining the free energy of subunit assembly (assembly of two dimers to form one tetramer) for the ligated state in question, rather than from measurements of ligand binding equilibria to the

tetramer directly. This is demonstrated in Figure 3 for binding of a single ligand at the α^1 site on the tetramer, but can be extended to any ligation state.

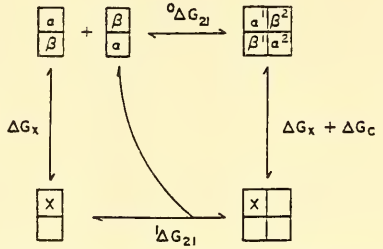


Figure 3 Thermodynamic coupling between subunit assembly (dimers to tetramers) and ligand binding at a single α subunit.

In Figure 3, $^0\Delta G_{21}$ is the free energy of forming an unligated tetramer (top) from two unligated dimers, and $^1\Delta G_{21}$ is the free energy to form a singly ligated tetramer (bottom) from a singly ligated dimer and unligated dimer. ΔG_X is the free energy difference between singly ligated and unligated dimers, and $(\Delta G_X + \Delta G_C)$ is the free energy difference between unligated and singly ligated tetramers. The sum around the cycle must be zero. Using the convention of taking reactions left to right and top to bottom,

$$^0\Delta G_{21} + \Delta G_X - \Delta G_C - {}^1\Delta G_{21} - \Delta G_X = 0,$$

and

$$\Delta G_C = {}^1\Delta G_{21} - {}^0\Delta G_{21}.$$

In general, the cooperative free energy for the formation of any ligated species, ij (see Figure 1), is given by the relationship

$$\Delta G_C = {}^i\Delta G_{2j} - {}^0\Delta G_{21},$$

where $^i\Delta G_{2j}$ is the free energy of formation of the tetramer, ij, from its constituent dimers.

Dr. Ackers found, in the three systems he studied (CN-met, MnIII, and MnII) that the ten ligation states of tetrameric hemoglobin exhibit only three distinct values of cooperative free energy. There are thus three principal energetic levels of the tetrameric molecule, after subtraction of the intrinsic (in the absence of site-site interaction) terms. In other words, for any intermediate molecular species, the free energy relative to that of the unligated (or fully ligated) molecule is a simple combination of the intrinsic free energy, $i\Delta G_X$, and one of the three cooperative free energies. The distribution of the ten ligation states of the tetramer among these

cooperative free energy states is shown in Figure 4 for the three systems studied. The three cooperativity states are equally spaced for the CN-met and Fe(II)/Mn(III) systems and almost equally spaced for the Mn(II)/Fe(II)-CO system. The total range of cooperative free energies is also approximately the same in the three systems (6 kcal, 7 kcal, 7.6 kcal) and is similar to tne cooperative free energy for oxygen binding to hemoglobin under the same conditions (6.3 kcal).

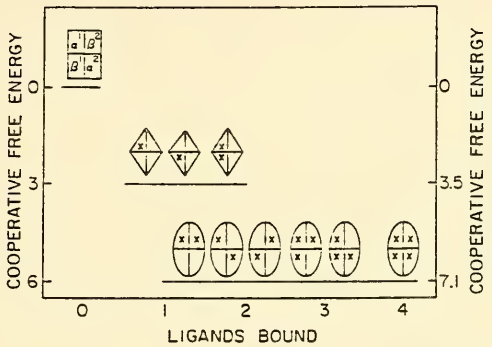


Figure 4 Cooperative free energy levels for tetramers in various ligation states. Scales pertain to the CN-met system (*left*) and the Fe(II)/Mn(III) system (*right*). *Ovals*, *diamonds*, and *rectangles* suggest three global states in the combinatorial mechanism.

These results suggest that there are three global states in the molecular mechanism of ligation (i.e., a three-state molecular switch mechanism rather than the heretofore accepted two-state mechanism). The discovery of three distinct cooperative free energies does not of necessity rule out the MWC two-state model. But Dr. Ackers has rigorously analyzed his data with respect to the thermodynamic requirements of the MWC model and has shown that the experimentally determined free energies are not compatible with the two-state model. The conclusion that at least three molecular states of the hemoglobin tetramer are required to account for the observed cooperativity of ligand binding implies that there are also at least three corresponding molecular structures. While no intermediate structures have been isolated, there is some tantalizing evidence from NMR studies of Drs. S. Miura and C. Ho that singly ligated molecules may have a different quaternary structure than that of either the unligated or fully ligated species.

Dr. Ackers points out that the results obtained from the non-oxygen systems he studied may not hold true for oxygen (i.e., it remains possible that oxygen binding proceeds via only two cooperativity states, as the MWC model predicts). However, in addition to the three ligand systems he studied extensively, Dr. Ackers has qualitative evidence for a similar pattern of energetic effects for CO or NO as ligands. Thus, every system studied to date shows evidence for more than two cooperativity states.

Dr. Michael Johnson of the University of Virginia (GM 28928) has also recently carried out an extensive analysis of oxygen binding data in order to test Dr. Ackers' proposed three-state switch model. He actually found that a four-state switch mechanism, rather than a three-state, was required to explain these data. He points out that his findings are not necessarily in disagreement with Dr. Ackers' data, since the third and fourth cooperativity states derived from his calculations are very close in energy (0.5 kcal); they might not have been resolved in the experiments carried out by Dr. Ackers. The significant point of Dr. Johnson's analysis is that it also supports a mechanism for the control of oxygen binding to hemoglobin that requires more than two states.

Clearly, additional investigation will be needed to fully resolve the molecular mechanism by which ligand binding to hemoglobin proceeds. The relationship of the new energetic findings to structural aspects of the mechanism will be an important area of future research. There have been attempts to crystallize intermediate-state tetramers. In all studies to date, the hemoglobin molecules have crystallized in either a deoxy-like or oxy-like structural form (there have been variations in subunit tertiary structure, but no new quaternary forms have been found). This may mean that additional cooperativity states fall among the structures usually called deoxy (T-state) or oxy (R-state), or that other quaternary structures do not crystallize nearly as readily as the two end-state forms. The identification of distinct intermediate ligation state structures would provide substantial support for a model with greater than two states. While it is often difficult to isolate reaction intermediates, Dr. Ackers is confident that these states exist and that they will eventually be isolated. In the meantime his elegant thermodynamic studies, which are now being extended using new cryogenic methods, provide the basis for an exciting new look at this important protein system.

Reference

Ackers, G.K. and Smith, F.R., Ann. Rev. Biophys. Chem., 1987. 16:583-609.

Straume, M. and Johnson, M.L., Biochemistry., 1988. 27:1302-1320.

Helen Sunshine

Helen R. Sunshine

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date May 4, 1989

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Structural Studies of Flavocytochrome b_2,
1 R01 GM 20530, Dr. F. Scott Mathews, Washington University, St. Louis

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS
Chief, Biophysics Section, BPS, NIGMS

Biological oxidation and electron transfer are metabolic processes fundamental to all living organisms. These processes are important in providing energy for cellular growth and maintenance of function. The enzymes that control and catalyze the chemical reactions involved in these processes must direct the energy released, often in the form of electrons derived from the metabolite, to other components in the cell. One important class of redox active enzymes are flavoproteins. These enzymes can abstract a pair of electrons from a substrate during a dehydrogenation reaction and pass them on to an electron acceptor protein. The latter is often a cytochrome, a small heme-containing electron storage protein. The electron transfer complex formed by the two proteins is usually transient in nature and difficult to observe, so one must rely on model building procedures to gain insight into the structure of the complex.

Recently Dr. Mathews and his colleagues at Washington University have successfully analyzed the structure of flavocytochrome b_2, a soluble mitochondrial tetrameric enzyme with each subunit containing both a flavin and a heme cofactor located in separate domains (Figure 1). The interaction between these two domains provides an experimentally determined model of an electron transfer complex. The enzyme carries out three distinct functions. First, the enzyme oxidizes the substrate, lactate, to form pyruvate and two electrons are taken up by the flavin ring. Second the electrons pass one at a time from the reduced flavin to the heme group. Finally, the flavocytochrome heme is reoxidized by cytochrome c, another mitochondrial protein which feeds the electrons into the cytochrome oxidase pathway. In this final step, a complex is formed between the flavocytochrome b_2 and cytochrome c molecules.

The flavin-binding domain contains about 400 amino acids and is composed of a barrel-like structure made up of 8 parallel β strands connected together by 8 α helices running in opposite directions to the strands. The flavin is located at the top of the barrel near the outside edge of the flavin-binding domain and faces the heme group contained in the cytochrome domain. The cytochrome domain is much smaller, containing about 100 amino acids and binds

the heme group in a cleft composed of 4 α helices. The heme and flavin rings are approximately coplanar and their edges are separated by about 9 A. There are no intervening main chain or side chain atoms to conduct the electrons and the arrangement of the cofactors is consistent with a tunneling mechanism for electron transfer.

The flavocytochrome b_2 molecule has its four subunits packed around a symmetry axis as shown in Figure 2. The four flavin-binding domains are packed tightly about this 4-fold rotation axis (which is non-crystallographic, i.e., not perfect symmetry) and the four cytochrome domains are on the outside rim of the molecule. Because of the packing arrangement of the molecule in the crystal lattice, two independent copies of each subunit are observed, i.e., two of the subunits are in one state and two in a different state. In the first state both domains are fixed in position and the chemical environment of the two interacting cofactors is clearly defined. In the second state only the flavin-binding domain is fixed in position while the cytochrome domain is quite mobile and not visible in the electron density map. The folded chains of the cytochrome domain probably exist in several different discrete orientations throughout the crystal, but are disordered on the average. This indicates that when the molecule is in solution the cytochrome domains are free to move with respect to the flavin-binding domain. The observed interaction between domains in the other subunit (which is in the first state) is truly representative of a transient electron transfer complex.

There is another interesting difference between the two observed states of the flavocytochrome b_2 molecule. In the first state, the cytochrome domain is fixed in space with a single ordered structure and the active site in the flavin binding domain is empty, containing only water molecules. In the second state, the cytochrome domain is mobile and the active site is occupied by a molecule of pyruvate, the product of the reaction. These results suggest that the presence of pyruvate in the active site and formation of the intramolecular electron transfer complex may be mutually exclusive. Inhibitor binding studies carried out in Dr. Mathews' laboratory show that the active site is accessible in both subunits, indicating a special mechanistic role for the bound product molecule.

The results described above have provided structural details on the relative orientation of the heme and flavin cofactors in the electron transfer complex as well as on how the cofactors interact with the protein atoms. The location of the product in the active site has also helped define the substrate specificity and catalytic mechanism of proton abstraction during the reaction. Studies are currently underway in Dr. Mathews' laboratory on a number of genetically engineered mutants of flavocytochrome b_2. The aim of these studies is to learn precisely how certain amino acids are involved in

the enzymatic activity and the electron transfer processes and possibly to alter the activities of the protein in a useful way. For example, several side chains in the active site, such as Tyr 274 and Tyr 143 have been altered and found to affect one or more of the three catalytic functions of the enzyme. By combining studies of the mutant enzyme in crystals and solution further insight into oxidation-reduction and electron transfer reactions will be obtained.

John C. Norvell

John C. Norvell, Ph.D.

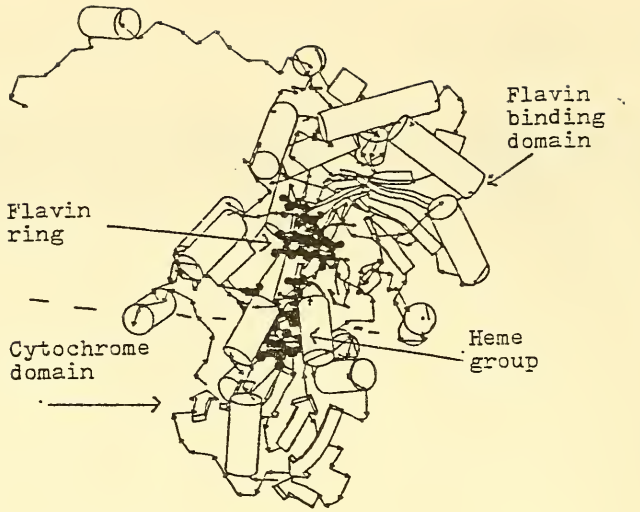


Figure 1. Diagram of one subunit of flavocytochrome b_2. The flavin-binding domain is above and the cytochrome domain is below. In this diagram cylinders represent α helices and arrows represent β strands.

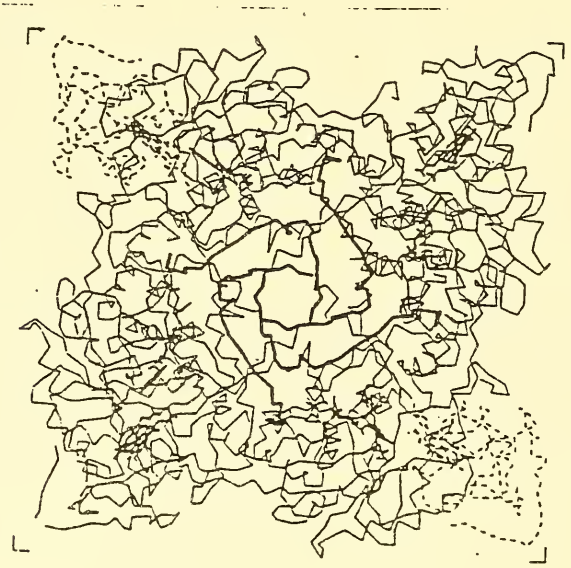


Figure 2. C_α tracing of the flavocytochrome b_2 tetramer viewed down the 4-fold axis. Dashed lines are the positionally disordered cytochrome domains. The hemes, flavins, and C-terminal tails are drawn in heavy lines.

Memorandum

Date August 7, 1989

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Structural Studies on Proteins Involved in Hemostasis, 5 R01 GM 33192-05, Dr. Brian F.P. Edwards

To Director, NIGMS
Through: Director, Biophysics and Physiological Sciences Program, NIGMS
Chief, Biophysics Section, BPS, NIGMS

Hemostasis is the process for arresting the flow of blood from an open vessel. It involves the delicate interplay among the blood vessel walls, platelets, and blood-clotting proteins. The long term goal of this project is to explain the function of important proteins in hemostasis by determining the x-ray crystallographic structures of these proteins alone and complexed with various ligands. Recently, Dr. Edwards and his coworkers have determined the three dimensional structure of two such proteins, thrombin and platelet factor 4.

Thrombin is the central enzyme in the regulation of hemostasis. It has a 49 residue A-chain which is linked by a disulfide bond to a 259 residue B-chain. At low concentrations it activates platelets and interacts with vascular endothelial cells; at higher concentrations it activates factors V, VIII, and XIII and converts the protein fibrinogen into fibrin, causing the blood to clot. When bound to thrombomodulin on cell surfaces, thrombin retards clotting by activating another protein, protein C, which shuts off the coagulation cascade. There is a great interest in developing clinically useful inhibitors of thrombin because uncontrolled clotting by this "killer enzyme" is a major factor in approximately 40% of all deaths.

Dr. Edwards and his coworkers (Dr. Philip Martin, Dr. Vinod Kumar, Mr. Raju Kunjummen) originally solved the structure of thrombin at 3.5 A by molecular replacement, i.e.,initiating the structure solution by using the segments of the B-chain that were homologous with other serine proteases such as chymotrypsin and elastase and using them as the search models for the molecular replacement computer programs. Recent work has extended the model to 2.4 A resolution. Although the structure of several mobile loops on the surface of thrombin remains to be determined, some features of the overall structure, including the location of 12 residues of the A-chain, are readily apparent (Fig.1). The structure is similar to chymotrypsin or elastase, in particular, the position of the catalytic serine is essentially unchanged, but the catalytic histidine in thrombin is rotated approximately 90° and the catalytic aspartate is shifted. Once the structure of the free enzyme is completed, Dr. Edwards will use this information to determine the structures of several modified thrombins, which have been crystallized with inhibitors bound.

Platelet factor 4 (PF4) is a protein that is released by activated platelets. It binds tightly to heparin, a sulfated polysaccharide that is a powerful anticoagulant, present on cell surfaces and in plasma. Platelet factor 4

interacts with the blood-clotting system primarily by neutralizing the effects of heparin and related molecules. Dr. Edwards and his coworkers (Mr. Robert St. Charles and Dr. Daniel Walz) have determined the x-ray crystallographic structure of PF4 at 3.0 A resolution. Platelet factor 4 is the first heparin-binding molecule, of which some 50 are known, whose structure has been determined to atomic resolution.

Bovine PF4 contains four identical subunits, each comprised of 85 residues. Each subunit has an extended loop and three strands of antiparallel β-sheet (shown as arrows in fig.2) on the inside and one α-helix on the outside (Fig. 2). It differs significantly from the structure predicted from an analysis of secondary structure and chemical modification experiment. The dimer of PF4, which is shown in fig. 2, has an extended six-stranded β-sheet formed by two subunits binding side by side. The two α-helices are arranged antiparallel to one another on the exterior of the extended α-sheet. The α-helices contain the four carboxy-terminal lysine residues that are primarily responsible for binding heparin. The tetramer of PF4 is formed by two dimers that are arranged back-to-back -- with β-sheets on the inside and α-helices on the outside. An unusual feature of the tetramer is the presence of potential salt bonds buried inside the oligomer between glutamic acid 43 of one subunit and the lysine 65 residue in the opposing subunit.

The PF4 structure has general significance beyond blood coagulation and the fact that it is the first structure of a heparin-binding protein that has been determined. Several recently-discovered proteins that are homologous to PF4 are involved in regulation of cell growth. Members of this class are induced by γ-interferon and platelet derived growth factor, constitutively overexpressed in transformed cell lines and strongly induced by Rous sarcoma virus in fibroblast cell lines. As a class, these proteins seem to be associated with cell growth and wound healing.

Experiments by other investigators using small angle neutron scattering and chemical modification suggest that heparin wraps around the outside of the PF4 tetramer and that four lysines in the carboxyl terminal helix of each subunit are major components of the binding site. The crystal structure shows that the helices are arranged in pairs. The heparin polysaccharide chain could bind either between the helices or across them at an angle. A calculation of the electrostatic field around PF4 suggests that heparin binds across the two helices.

In these important studies, Dr. Edwards and his colleagues have made substantial progress in understanding the molecular basis of hemostasis.

John C. Norvell

John C. Norvell, Ph.D.

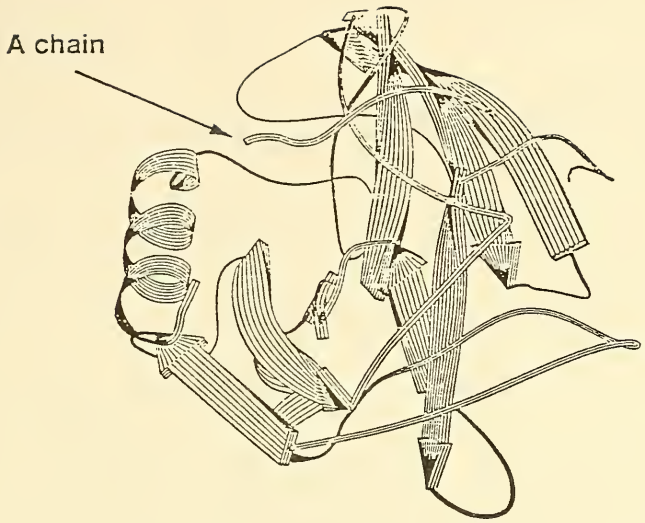


Figure 1: A ribbon diagram of thrombin showing the A-chain on the side opposite the active site. Several small gaps in the current tracing of the B-chain have been automatically closed by the drawing program.

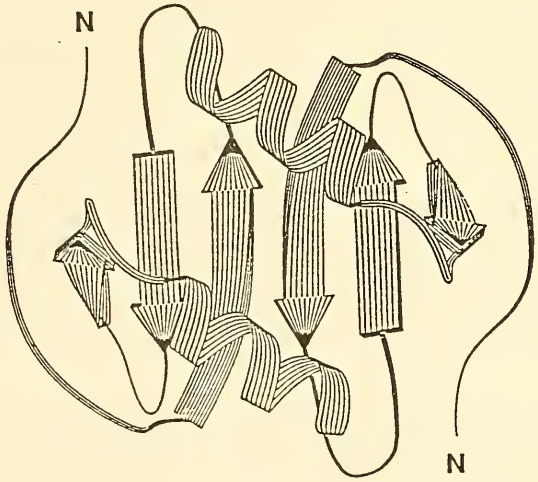


Figure 2: A ribbon diagram of the PF4 "dimer" showing the two antiparallel helices that form the heparin binding site.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 21, 1988

From Program Administrator, Biophysics Section, BPS, NIGMS

Subject RESEARCH HIGHLIGHT: "Bioanalytical Microseparations, Detection and Perfusion"
Ewing, Andrew G., 5 R01 GM37621-03

To Director, NIGMS
Through: Director, BPS, NIGMS See for me
Through: Chief, Biophysics Section, BPS, NIGMS

Analytical chemists daily deal with a number of realities concerning any qualitative or quantitative analysis performed, namely the specificity and sensitivity of the reactions. Since there are only a relatively small number of chemical reactions which are unique for specific compounds, the analyst is almost always faced with the problem of knowing what is being measured in any given reaction. In addition, many times reactions are negative due to the lack of sensitivity of the test in question. In the past it was necessary to use labor intensive methods to isolate and purify the compounds of interest in order to measure the amounts present in the original samples.

The advent of chromatographic and electrophoretic techniques has greatly simplified the problems of specificity. These techniques permit the separation of substances within a mixture before their presence is determined by specific detector systems. Most detectors, however, do not detect all of the substances separated: i.e., an ultraviolet light detector will discern only those substances which absorb ultraviolet light. Substances which do not absorb ultraviolet light will not appear on the chromatogram even though the substance passed through the detector.

A chromatogram normally shows a series of peaks and valleys. These valleys, however, are not to be interpreted as void spaces lacking chemical entities. These void spaces may in fact contain substances which have passed through the system without being detected. Good universal detectors, which are responsive to all chemical substances, are needed before it will be possible to account for all the substances present in a given sample. Conversely, in the absence of a good universal detector, the purity of the isolated peaks cannot be unequivically established. Within these limitations, as the amounts (weight and/or volume) of biological samples to be collected and analyzed decrease, the problems associated with their analyses increase.

Dr. Andrew G. Ewing, Assistant Professor of Chemistry at the College of Science of the Pennsylvania State University, with the support of the National Institute of General Medical Science (5 R01 GM37621-03), has been working to

perfect the detection of neurotransmitters at a cellular level. The problem faced by him is that there are several neurotransmitters, and no one method which will completely suffice for their separation.

Dr. Ewing selected the method of Capillary Zone Electrophoresis because he felt this technique offered the opportunity for the greatest reduction in sample size as he scaled down the size of the capillary system of this electrophoretic apparatus. Furthermore, it offered the potential for the greatest improvement in separation efficiency because the smaller the capillary the better its heat dissipation properties during its operation. This improvement in temperature control would decrease the dispersion of the solute due to diffusion. In his most recent work, Dr. Ewing has reduced the capillary diameter to 26 microns from 75 microns in order to decrease his sample size to 0.23 nL.

In ordinary electrophoresis cations migrate to the cathode and anions migrate to the anode. In capillary electrophoresis the effects of the zeta potential are more pronounced. The resistance offered by the capillary to the movement of the solute under the applied potential causes the solvent to move and establishes an electroosmotic flow. This flow of the solvent tends to carry all migrating solutes in the direction of the anode. The anions flow ahead of the neutral compounds and the latter ahead of the cations. This peculiarity of capillary electrophoresis permits the detection of all species with a single detector placed near the anode.

Dr. Ewing used an electrochemical detector which he had modified and adapted to use with his smaller capillary systems in order to greatly increase his detector sensitivity. Other detectors, such as colorimeters might have been used had the increased sensitivity not been required. In order to extend the advantages of Capillary Zone electrophoresis to the neutral compounds a technique known as Micellar Electrokinetic Capillary Chromatography was incorporated into the protocol.

Micellar Electrokinetic Capillary Chromatography is a technique whereby the electroosmotic flow, a phenomena normally associated with capillary electrophoresis, is being utilized for the separation of neutral compounds. This is accomplished by adding surfactants to the electrophoretic buffer system to establish a micelle system of two immiscible phases. The neutral compounds can then distribute themselves between the two phases and achieve counter current chromatographic separation as the fluids move through the capillary system.

Dr. Ewing has been able to separate neutrotransmitters at a level of 20 femtomoles from a sample size of 0.23 nL with an efficiency of 400,000 theoretical plates. His current experimental setup is a 26 micron capillary system running at 20 kV. He anticipates being able to prepare capillary systems with a diameter of 9 microns. He feels this improvement will allow him to use a sample size of 18 pL and detect attomole and even subattomole quantities (several thousand molecules) of the neurotransmitters. At this point Dr. Ewing feels he will be ready to assay the contents of single nerve cells.

The figures below show the separation of neutrotransmitters with and without micellar formation: 1) Norephinephrine, 2) Epinephrine, 3) 3,4-Dihydroxybenzylamine, 4) Dopamine, 5) L-DOPA, 6) Catecol, 7) 4-Methylcatechol.

Américo Rivera, Jr., Ph.D.

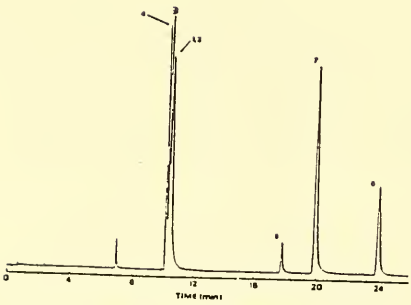


Figure 1: Capillary Zone Electrophoresis

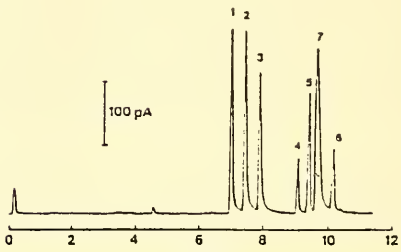


Figure 2: Micellar Electrokinetic Capillary Chromatography

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 3, 1988

From Program Administrator, BPS

Subject RESEARCH HIGHLIGHT; "Liquid Chromatography of Biological Substances"
2 R01 GM 20993-14 (Horvath, C.)

To

Director, NIGMS
Through: Director, BPS
Acting Chief, Biophysics Section

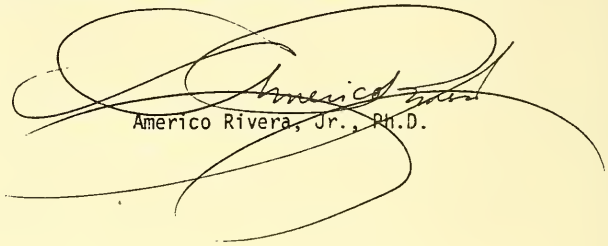
The extraordinary impact of modern liquid chromatograhy on the life sciences and medicine for the last decade affirms the observation of M.S. Tswett, the inventor of chromatography, that every scientific advance is an advance in method. There is every reason to believe that the technique will continue to develop and/or diversify at an ever accelerating pace. As a result, the exploitation of the full potential of this technique on a broad scale is yet to come. During the last decade extraordinary advances have been made in liquid chromatography. The system has become more flexible and has rivaled gas chromatography in every parameter except for time. Now even this limitation has been resolved. High performance liquid chromatograhy is every bit as versatile as gas chromatography for compounds which cannot be analyzed by this latter method.

This major breakthrough was made recently in the laboratory of Dr. Csaba G. Horvath, Professor of Chemical Engineering at Yale University with the support of the National Institute of General Medical Sciences (GM 20993-14). Dr. Horvath has established himself as an international authority in both the theoretical and practical aspects of reverse phase liquid chromatography of biological compounds and substances. His recent discovery that liquid chromatography columns can be developed in seconds by elevating their temperature to 80 degrees now makes liquid chromatography equivalent to gas chromatography techniques in speed and sensitivity. However, Dr. Horvath and his laboratory are best known for their long standing interest in the development of support media and how changes in the stationary phase effect the separation of different compounds and families of compounds.

In addition to Dr. Horvath's major contributions to the advancement of the science of chromatography, he also has been interested in the theoretical aspects of chromatographic processes. There has been a great need to understand the retention mechanism of macromolecules. With his extensive experience and database of the reverse phase chromatographic behavior of a large number of compounds to different support media, Dr. Horvath has been able to develop a series of mathematical expressions which yield acceptable and satisfactory expressions of a logarithmic retention factor. These expressions take into account the physical and chemical nature of the stationary phase, the mobile (solvent) phase and the solute distributed between the two phases.

Dr. Horvath is interested in expanding these equations to allow for the prediction of the retention factor for proteins. However, the current state of the art requires information concerning the presence of key chemical structures. Such information is not available for most of the proteins found in biological solutions. Nevertheless, efforts continue in his laboratory to extend these models as much as possible to include retention predictions of more complicated compounds. Dr. Horvath hopes that the models will be instrumental in helping design new separation systems which would separate compounds which are now intractable. Furthermore, he feels that these systems will help in the design of preparative processes for specific substances.

Dr. Horvath has also made a number of contributions to the isolation and characterization of glycans, and proteins. Dr. Horvath's work has had a major economic impact in the pharmaceutical and biotechnology industries where these chromatographic techniques are crucial for the analysis and separation of specific substances.



Americo Rivera, Jr., Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 19, 1988

From Program Administrator, BPS

Subject RESEARCH HIGHLIGHT: "Cell Characterization with Ultrasound" 5 R01 GM 30419-06 (Apfel, R.)

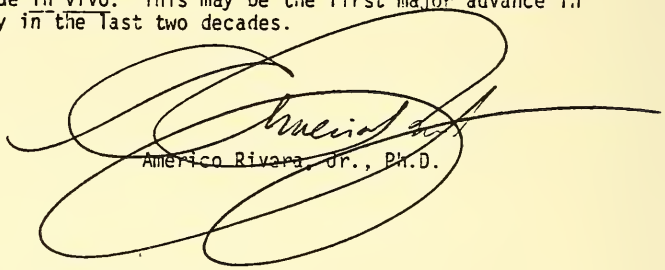
To Director, NIGMS
Through: Director, BPS
Acting Chief, Biophysics Section

For several decades the primary interest of the ultrasound research community was the improvement of sonic reflective images and the determination of the mechanism by which the harmful effects of ultrasound may take place. The problems associated with the extraction of medical information from the sonic scatterers were largely ignored. There had been some earlier work with biological fluids and solutions that revealed a non-linear component, "B/A," of ultrasound absorption which was directly proportional to the concentration of proteins in these solutions.

Professor Robert E. Apfel, a medical engineer at Yale University, has been pursuing three projects involving the characterization of biological material with acoustic waves: 1) cell characterization using acoustic scattering; 2) cell deformation produced by acoustic radiation forces; and 3) tissue characterization by precise measurement of the acoustic nonlinearity parameter, "B/A." Considerable progress has been made on the first and last projects, so much so that Dr. Apfel proposes to design two novel instruments which will be used to characterize tissue properties using ultrasonic energy.

The first instrument will be a cell sorter which will measure cell density and compressibility using scattered radiation from high frequency acoustic irradiation of cells. Long wavelengths will be used so that cell size will not be a factor in the data collected. The instrument will measure cell volume with an electrozone sensing element (a two electrode sensor which uses the Coulter principle to estimate particle volume). The three variables, compressibility, density, and volume, of each cell will be the parameters used to sort cells into different populations. Experiments have shown that different cell populations in human blood buffy coat can be resolved. It was also possible to independently characterize the mean density and compressibility of red blood cells and granulocytes. He is encouraged by these data and the early reports in the literature of the sonographic identification of tumor tissue. Dr. Apfel proposes to measure these mechanical properties on a variety of heterogeneous cell populations. He feels that the characteristics being measured will distinguish between normal and diseased states and may even be used to sort cancer cells from normal cell populations. This is a major advance in the field of cell sorting. Cell biologists have been trying for the last 5 years to add ultrasound as one of the modalities with which to sort for physical cell properties.

In addition to the above work, Dr. Apfel has developed a theoretical model which suggests that the percentage of water, protein, and fat can be determined from ultrasonic measurements. This group has obtained preliminary data to show that ultrasonic radiation model calculations for the percentage of water, protein, and fat in a sample are in close agreement with experimental data. Encouraged by these results, Dr. Apfel also proposes to build a second instrument to measure the non-linear parameter, "B/A," of tissues and biological samples. They propose to develop an instrument to make these measurements and to define and refine their algorithms to permit the calculation for the percentage of water, protein, and fat in a biological sample. Furthermore, Dr. Apfel feels that under appropriate conditions the measurements of the the non-linear parameter, "B/A," may be used as an imaging modality which may combine tissue composition modeling with special information regarding tissue in vivo. This may be the first major advance in ultrasound imaging modality in the last two decades.



Americo Rivera, Jr., Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 30, 1989

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Molecular Pathology of LPS-Induced Shock and DIC, 5 R01 GM28485-08 (Ulevitch, R.), Scripps Clinic and Research Foundation

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program, NIGMS 4/10/89

Gram-negative sepsis is a frequent clinical complication of major injury. A significant number of patients with gram negative sepsis develop a syndrome known as multiple organ failure (MOF) that is often fatal. MOF is often preceded by fever, hypotension, leukopenia and evidence of disseminated intravascular coagulation (DIC). Because many different types of gram-negative bacteria produce this common clinical picture, the search for a toxin ubiquitous among all gram-negative organisms was the focus of numerous studies during the early 1950's. It is now well accepted that the endotoxin or lipopolysaccharide (LPS) from the outer membrane of gram-negative bacteria initiates the cascade of events that results in shock, DIC, and MOF. While LPS has been studied extensively during the past twenty years, it has only been recently that a basis for understanding the molecular mechanisms of LPS-induced injury has emerged. Key in achieving this understanding is the recognition that the macrophage plays a central role in the host response to LPS.

Dr. Ulevitch's laboratory has provided much of the evidence to support the role of the macrophage in LPS-induced injury. Results from his research indicate that LPS has the capacity to stimulate macrophages to release multiple injurious substances. These include oxygenation products of arachidonic acid, toxic O_2 products, degradative enzymes, tumor necrosis factor (TNF) and several other monokines with a wide range of activities.

Of the many macrophage products that play an important role in the regulation of the host response to LPS, TNF has been the one mediator demonstrated to have the potential to initiate the full range of pathophysiologic changes observed after an intravenous LPS injection. For example, infusion of TNF (purified from rabbit peritoneal macrophages) results in fever, altered perfusion, lactic acidosis, intravascular coagulation, and increased systemic and pulmonary permeability. In addition, TNF causes severe organ failure involving the lungs, liver, kidneys, and gastrointestinal tract. In regards to circulating levels of TNF, increased TNF levels have been observed in the blood of patients dying of sepsis but not in survivors.

In more recent investigations, Dr. Ulevitch has studied the effects of administration of rabbit anti-TNF globulin on LPS-induced injury in a rabbit model of endotoxin shock. Administration of anti-TNF globulin prior to

intravenous injection of LPS had several marked effects. The gradual fall in blood pressure that occurs during a six hour period following intravenous LPS injection was markedly reduced. Liver and kidney tissues removed at the time the animals were sacrificed showed a marked reduction in cellular necrosis and fibrin deposition in anti-TNF treated animals when compared to non-immune globulin treated rabbits. In contrast, the fall in blood leukocyte numbers and the increase in core temperature produced by LPS was not altered by pretreatment with anti-TNF globulin. Most importantly, the administration of anti-TNF prevented the death of LPS-treated rabbits in a situation that would otherwise result in the death of all LPS-treated rabbits.

An unexpected finding of these studies was that following a single injection of LPS, the rabbits developed a refractory period during which time they did not produce detectable TNF in the blood after a second LPS challenge. Dr. Ulevitch put forth a number of possibilities to explain this phenomenon. He suggested the possibility that inhibitors of TNF were responsible for the refractory state which occurs with the second LPS challenge. However, examination of post-LPS serum failed to yield evidence for the presence of inhibitors of the TNF assay or for substances that would inhibit the induction of TNF. He also tested the possibility that TNF was down regulating itself by infusing human recombinant TNF into the rabbits, and after several hours injected the animals with LPS. In these animals there was the expected level of endogenously produced TNF, suggesting that TNF itself or products induced by TNF are not involved in the down regulation phenomenon. A third possibility was that the down regulation occurs as a result of a direct effect of LPS on the macrophage. To pursue this Dr. Ulevitch initiated studies with a murine macrophage cell line and with rabbit peritoneal exudate macrophages. In both cases, exposure of these cells to LPS (0.1 - 10 ng/ml) for a minimum of six hours led to a substantial reduction in TNF production when the cells were rechallenged with a maximum stimulatory dose of LPS (100 ng). Dr. Ulevitch suggests that the down regulation involves a post-transcriptional block, since in separate studies TNF mRNA was found to be induced in the LPS pretreated cells; while the TNF mRNA was apparently not translated.

Dr. Ulevitch is extending the above studies and investigating the regulation of TNF gene expression. He is focusing on the TNF-alpha gene. The goal of this work is to define the kinetics of TNF-alpha gene induction by LPS in macrophages, to relate this induction to expression of the TNF-alpha protein, and to evaluate whether the transcriptional stabilization of mRNA contributes to the inductive process. It is anticipated that these studies will provide information about the signal transduction pathways that regulate LPS-induced TNF-alpha production at the RNA and protein levels; thus, providing a better understanding of the regulation of TNF-alpha production in vivo. A determination of the molecular events associated with TNF production should help in the development of strategies for the prevention and treatment of gram-negative sepsis.

Yvonne T. Maddox

Yvonne T. Maddox, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date June 27, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Tumour Necrosis Factor (TNF) Regulation of Lipoprotein Lipase (LPL) mRNA in Adipocytes, 5 R01 GM32892-03 (Pekela, P.), East Carolina University

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program, NIGMS

One of the clinical signs of chronic infections or tumours is the presence of a catabolic state that can proceed to cachexia (a profound and marked condition of general ill health and malnutrition), shock, and death. The biochemical basis for this phenomenon is not understood. In order to gain insight into the mechanism of this process, Dr. Pekela and colleagues have been studying as a model system the hyperlipidemia that occurs with infection or endotoxemia. These investigators were the first to demonstrate that the hyperlipidemia associated with endotoxemia is mediated by an endotoxin-induced macrophage product (tumour necrosis factor, TNF), which markedly suppresses the activity of lipoprotein lipase (LPL). LPL is the enzyme which catalyzes the hydrolysis of lipoprotein-derived triacylglycerol, providing fatty acids for storage or catabolism. Its activity in adipose tissue allows the accumulation of fatty acids, which are re-esterified in the fat cells (adipocytes) to form stored triacylglycerol. Lipoprotein lipase is a short-lived enzyme with a half life of approximately 30 minutes and is subject to regulation by both nutritional and hormonal signals. A marked loss of enzyme activity occurs upon fasting. Also in various disease states such as chronic infections and malignancies, LPL activity is diminished. This leads to derrangements in lipid metabolism resulting in severe hypertriglyceridaemia and an overall catabolic state.

Regulation of the onset of this catabolic state, as well as the suppression of LPL activity, have been attributed in general to the monokines, which are secondary proteins released from phagocytic cells. When these proteins interact with specific plasma-membrane receptors on target tissues, they regulate various metabolic pathways. In adipocytes, for example, three monokines, i.e. TNF, interleukin-1 (IL-1), and gamma interferon, have been shown to attenuate the activity of LPL. Furthermore, TNF has been implicated in mediating cachexia.

Dr. Pekela has demonstrated that TNF suppresses LPL activity in 3T3-L1 adipocytes by inhibiting the synthesis of this enzyme. Several studies have been conducted to support this hypothesis. In the first set of studies a dose-response experiment was performed to test the effect of TNF on the activity of LPL. The results indicated that half-maximal suppression of the lipase activity occurred when cells were exposed to 3 ng of TNF. A maximal suppression of approximately 90% of enzyme activity was observed when 30.5 ng of TNF was added to the cells.

A reduction in lipase activity was apparent within one hour after addition of the monokine, while a maximal loss of activity was observed after 17 hours of incubation.

The experiment which provided strong evidence that TNF inhibited LPL synthesis involved exposing adipocytes to [35S]-methionine. Following [35S]-methionine labelling of the cell proteins, the [35S]-methionine labelled lipoprotein lipase was immunoprecipitated, analyzed by SDS-polyacrylamide gel electrophoresis and quantified with a densitometer. When the adipocytes were incubated with TNF the incorporation of [35S]-methionine into immunoprecipitable LPL was inhibited.

Until recently there has been no direct information on the turnover of LPL mRNA. However, the rapid decrease of LPL synthesis in the 3T3-L1 adipocytes exposed to TNF, coupled with the short half-life for enzyme activity, suggests that the mRNA is rapidly turned over. Dr. Pekela's approach to this hypotheses has been to determine by Northern analysis the molecular size of mRNA coding for LPL by using RNA isolated from both undifferentiated and fully differentiated 3T3-L1 cells as well as from mouse heart (as a comparison). Previous work by other investigators had demonstrated the presence of two major species of mRNA for LPL in mouse heart. Hybridizations were carried out in the presence of a nick-translated LPL cDNA probe. After autoradiography, it was possible to identify two major species of LPL mRNA in 3T3-L1 cells. He also found two major species of mRNA for LPL in mouse heart, and these species were identical with that found in the 3T3-L1 cells. The molecular size of these mRNA species as determined by using calibrated RNA standards was 3.7 kb and 3.9 kb.

In order to characterize the effect of TNF on amounts of mRNA coding for LPL, total RNA was prepared from 3T3-L1 cells of the same passage and plated at the same time. TNF (1.5nM) was added to undifferentiated and fully differentiated cells. From each preparation, 20 mg of RNA was used for densitometric measurements from northern analysis. In this experiment a 10-fold increase of LPL mRNA upon differentiation was observed, with an increase in LPL activity (100-fold). Exposure of the cells to TNF decreased the amount of LPL mRNA by nearly 60% and the LPL activity was depressed by 76%, evidence that the molecular basis for the loss of LPL activity is a significant decrease in LPL mRNA available for translation. After 48 hours of continuous exposure to TNF, the amounts of LPL mRNA returned to 85% of control and enzyme activity to 80% of control. This provides further support that the monokine is a potent, but reversible, regulator of adipose-tissue metabolism.

A single dose of TNF added at the induction of differentiation of 3T3-L1 cells did not, however, entirely stop the differentiation process. Of course, the results reported here on the effects of TNF on mRNA may differ upon exposure of the cells to multiple doses of TNF. This requires further investigation, together with experiments on the effects of TNF on the transcriptional process.

These are new and exciting studies which seem to demonstrate that TNF reversibly down-regulates LPL mRNA in fully differentiated 3T3-L1 adipocytes. Also cells induced to differentiate in the presence of TNF are likely to

exhibit a delayed time course for development of the adipocyte phenotype, as judged by the attenuation of the normal increase in LPL mRNA that occurs with differentiation. Furthermore, these investigations by Dr. Pekela suggest that pools of mRNA which code for the lipase are depleted rapidly in the presence of TNF. Whatever the mechanism of suppression of LPL activity, TNF interacts with adipose cells in a specific manner which halts the synthesis of the lipase. Hence these lipase-deficient cells would be compromised and unable to obtain fatty acids from triglycerides due to hampered triglyceride storage. More studies are needed to determine if the observed effects are direct effects of TNF or whether they are due to other hormones which are released in response to TNF.

Yvonne T. Maddox, Ph.D.

Memorandum

Date September 6, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Mechanisms of Neutrophil Dysfunction in Trauma, 5 R01 GM31754-06 (Solomkin, J.), University of Cincinnati

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program, NIGMS

The central role of neutrophils (PMNs) in the acute inflammatory process is well established and complement-mediated neutrophil activation (CMNA) represents one facet of the acute inflammatory process. The interaction of activated complement components, most notably C5a, with circulating PMNs initiates a series of PMN activation responses. These activation responses, necessary for effective microbicidal activity, include increased adherence and aggregation, chemotaxis, lysosomal enzyme release, arachidonic acid metabolite release, and superoxide production. In addition, CMNA causes several postactivation changes that up- or down- regulate subsequent neutrophil function. These modulating changes include chemotactic desensitization and enhanced oxidative responsiveness and are believed to further enhance PMN proinflammatary activity. However, the manner in which CMNA induces these postactivation changes is poorly understood.

Laboratory experiments have shown that CMNA and its resultant by-products cause endothelial cell injury. In addition, in animal models that utilize the intravenous infusion of activated complement, microvascular lung injury has been demonstrated. The clinical correlation of these findings is that CMNA has been implicated as a pathogenic mechanism contributing toward the development of adult respiratory distress syndrome (ARDS). This form of microvascular lung injury occurs in several clinical conditions associated with CMNA, including severe sepsis, multiple trauma, and cardiopulmonary bypass.

Dr. Solomkin and his colleagues have been seeking to define the PMN's postactivation state following CMNA. He has focused his efforts on the changes in cell surface receptor expression and oxidative activity that occur in neutophils after exposure to zymosan-activated serum (an *in vitro* model for CMNA; zymosan serves as a potent source of C5a). Utilizing this *in vitro* model, Dr. Solomkin has characterized the effects of CMNA on human PMN superoxide (SO) production and N-formyl-methionyl-leucyl-phenylalanine (FMLP) receptor status. Superoxide (oxidation products) generation is associated with endothelial cell injury (vascular damage) and FMLP is a chemotactic factor. PMN superoxide production was measured as the superoxide dismutase inhibitable reduction of ferricytochrome c in a continuous spectrophotometric assay. A radioligand saturation binding assay was employed for assessing FMLP receptor status. One of Dr. Solomkin's most recent findings, while employing

this model of zymosan-activated serum (ZAS), has been to demonstrate that CMNA was associated with a 132% increase in FMLP-induced SO generation and a 110% increase in FMLP receptor expression in neutrophils from normal subjects.

Thus, CMNA was associated with an enhanced PMN oxidative activity in response to FMLP and with an increased expression of FMLP receptors on the PMN's surface membrane without producing changes in receptor affinity. Steroid pretreatment of the PMNs with methylprednisolone (MPSS) prevented the increased expression of FMLP receptors. The FMLP-induced PMN SO production was also prevented. These results have caused Dr. Solomkin to hypothesize that the enhanced oxidative responsiveness following CMNA is due, in part, to the mobilization and resulting increased expression of PMN surface membrane receptors. Further supporting this hypothesis is the observation that MPSS pretreatment blocks both the increased receptor expression and the enhanced oxidative responsiveness associated with zymosan exposure. In separate studies, Dr. Solomkin has also studied PMNs from patients with sepsis. These cells exhibited a nearly twofold increase in FMLP-induced superoxide production when compared to control cells that had not been challenged with ZAS. PMNs from septic patients were also characterized by a twofold increase in the number of FMLP receptors expressed on their surface.

Dr. Solomkin believes that his work (1) and that of Fletcher and Gallin (2) provide preliminary evidence that limited degranulation of specific granules, as C5a exposure is known to produce, is the mechanism whereby an intracellular pool of FMLP receptors is translocated to the PMN's surface. The modulating role of steroids such as MPSS might therefore work to prevent FMLP receptor mobilization and the consequent enhanced oxidative responsiveness associated with CMNA. Also, there is the possibility that CMNA enhances FMLP-induced oxidative responsiveness via a nonspecific membrane process unrelated to receptor mobilization. Dr. Solomkin argues against this possibility since FMLP receptor affinity was unchanged.

These investigations have potential clinical significance since blocking these CMNA-associated events could modulate the potential damaging effects of the increased PMN oxidative metabolism. However, it remains to be seen whether methylprednisolone pretreatment can prevent or ameliorate complement-mediated acute microvascular injury in the in vivo situation.

Yvonne T. Maddox, Ph.D.

REFERENCES:

1. Solomkin, J.S., Cotta, L.A., Ogle, J.D., et al.: Complement-induced expression of cryptic receptors on the neutrophil surface: A mechanism for regulation of acute inflammation in trauma. Surgery 96: 336, 1984.

2. Fletcher, M.P., and Gallin, J.I.: Degranulating stimuli increase the availability of receptors on human neutrophils for the chemoattractant f-met-leu-phe. J. Immunol. 124: 1585, 1980.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 7, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Lung Injury by Oxygen Metabolites, 5 R37 GM29507-07 (Ward, Peter), University of Michigan, Ann Arbor

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program 12/2/88

The exposure of neutrophils to inflammatory mediators, such as chemotactic factors, results in a respiratory burst which is characterized by increased oxygen uptake, stimulation of the hexose monophosphate shunt cycle and generation of oxygen free radicals. It is well-recognized that the respiratory burst of neutrophils is essential for bactericidal activity. This respiratory burst results from the activation of an adenine dinucleotide-requiring oxidase in the plasma membrane that catalyzes the production of superoxide from oxygen using NADPH as the electron donor. Superoxide (O_2^-) production leads to the formation of other oxidizing species, such as hydrogen peroxide (H_2O_2), hydroxyl radical ($HO^\bullet$), and singlet oxygen ($^1O_2$).

There are many reports suggesting an important biological role for these oxygen-derived free radicals in the induction of tissue damage. For example, involvement of these highly toxic oxygen radicals has been demonstrated in glomerulonephritis, lung microvascular injury secondary to skin burns, the adult respiratory distress syndrome and in lung injury in experimental animals following the infusion of endotoxin or complement-activated plasma.

Dr. Peter Ward has been focusing on mediators responsible for lung injury and has established a role for oxygen metabolites in the pathophysiology of the lung. Recently, his investigative group has provided experimental in vitro evidence which suggests that neutrophil-derived oxygen radicals may impair the function of the oxidant-producing neutrophil itself. The damage of neutrophils due to superoxide and hydrogen peroxide production may result in partial or complete loss of cellular function and can even cause cell death.

In order to gain insight into the possible mechanisms that could explain loss of neutrophil function resulting from this autoxidation, Dr. Ward and his colleagues investigated the effects of various antioxidants on the chemotactic responsiveness of peritoneal neutrophils from rats. These neutrophils were assessed for chemotactic migration to several different chemotactic peptides, including insulin-activated rat serum (contains C5a), N-formyl-methionyl-leucyl-phenylalanine and N-formyl-norleucyl-leucyl-phenylalanine. All three stimuli induced cellular migration. In order to establish a role for superoxides in the migration of neutrophils, the effects of two anti-oxidant enzymes, catalase and superoxide dismutase (SOD), which are specific for H_2O_2 and O_2^-, respectively, were also investigated. When catalase was present in

the cell compartment of the chemotaxis chamber, there was enhancement of cellular migration with each of the three chemotactic stimuli, whereas the addition of SOD had little effect. Two scavengers of hydroxyl radicals, namely mannitol and dimethylsulfoxide (DMSO) and the iron chelator, deferoxamine, also significantly enhanced the chemotactic responses of the neutrophils. If the deferoxamine was iron-saturated prior to addition to the cells, there was no enhancement of chemotaxis. These results have prompted Dr. Ward to suggest that hydroxyl radical is the most likely oxidant involved in neutrophil autoxidation. This reasoning stems from his prior observations on the role of iron in the formation of hydroxyl radicals.

Dr. Ward demonstrated that HO• generation from neutrophils requires extracellular iron and explained the role of iron to exist in one of two ways. One is the Fenton reaction:

$$Fe^{3+} + O_2^- \longrightarrow Fe^{2+} + O_2$$
$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH^- + HO^\bullet$$

The second possibility is the formation of perferryl ion:

$$Fe^{2+} + O_2 \longrightarrow Fe^{2+}O_2 \text{ (perferryl ion)} \rightleftharpoons Fe^{3+}O_2^-$$

The perferryl ion, an activated dioxygen-linked ferrous ion, is a transition metal complex, and iron in this form can act as a redox agent. In addition, the dioxygen in the form of perferryl ion has the characteristics of a free radical. There is also evidence that the perferryl ion can enter into reactions resulting in lipid peroxidation. Although the evidence is indirect, lipid peroxidation may be an important event linked to cell death.

It has been suggested that the enhanced chemotaxis, observed with the agents used in these studies, is the result of protective interventions that spare chemotactically activated neutrophils from autotoxicity caused by the toxic oxygen radicals. It is unlikely that the protective interventions are due to interference with neutrophil activation, since rat neutrophils that are stimulated with phorbol myristate acetate (a potent cell activator for the generation of oxygen radicals) produced as much O_2^- in the presence or in the absence of deferoxamine or DMSO.

These data suggest that catalase, hydroxyl radical scavengers, and iron chelators protect chemotactically stimulated neutrophils from autotoxicity caused by neutrophil-derived hydrogen peroxide and hydroxyl radical. The importance of Dr. Ward's observations lies with the essential role of the neutrophil in microbial resistance and acute inflammation and hence, the need to protect these cells from self-destruction.

Yvonne T. Maddox, Ph.D.

Memorandum

Date November 25, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Pathogenesis of Pseudomonas in Thermal Injury, 5 R01 GM24028-11 (Saelinger, C.), University of Cincinnati

To Director, NIGMS

Through: Director, BPS Program, NIGMS

Infections by Pseudomonas aeruginosa have been the leading cause of mortality and morbidity in burn patients for many years. Antibiotic therapy for these infections has not been as successful as it was once hoped, owing, in large measure, to the innate capacity for these microorganisms to develop resistance to a wide variety of antimicrobial agents. An additional reason for the severity of these infections is the fact that Pseudomonas produces a wide variety of factors which are reported to enhance virulence. Most research efforts have focused on two of these virulence-associated factors: exotoxin A and the proteolytic enzymes, alkaline protease and elastase. Although numerous investigations have attempted to determine the relationships between the elaboration of the particular virulence-associated factors and Pseudomonas infection, the role that these factors play in vivo is not clear. Both cell-associated factors and extracellular products contribute to the organism's ability to establish an infection.

Dr. Saelinger has been investigating the mechanisms of Pseudomonas aeruginosa exotoxin A (PE) internalization in the cells of animals following thermal injury. PE is a nonglycosylated polypeptide, bacterial toxin which inhibits protein synthesis in mammalian cells by catalysing the transfer of ADP from nicotinamide-adenine dinucleotide (NAD) to elongation factor 2 (EF-2). Because EF-2 is a cytoplasmic protein, PE must gain entrance to the host cell cytoplasm to reach its substrate. Therefore, to express biological activity, PE must bind to cells via specific surface receptors and be internalized and its enzymatically active form must reach EF-2 in the cytoplasm. Hence the events that lead to PE inhibition of protein synthesis can be divided as follows: (i) binding to cell surface; (ii) internalization; (iii) intracellular trafficking; (iv) conversion of toxin to an enzymatically active form; (v) escape of the enzymatically active form into the cytosol and (vi) ADP ribosylation of EF-2 and subsequent inhibition of host cell protein synthesis.

Dr. Saelinger's group has shown that the toxin enters sensitive mammalian cells by receptor-mediated endocytosis. Endocytosis is the process whereby cells internalize substances from the external medium. It is divided into

phagocytosis and pinocytosis. Phagocytosis involves the internalization of particulate matter and large macromolecules (greater than or equal to 0.3 um) and is usually associated with macrophages, monocytes and neutrophils. Pinocytosis is the internalization of smaller macromolecules and solutes (less than 0.3 um). Thus, the internalization of PE is a pinocytic event.

In more recent studies, Dr. Saelinger and colleagues have conducted experiments designed to determine the site (organelle) from which the enzymatically active form of PE enters the cytoplasm. These studies involved a combined morphological and biochemical approach. Using ^{125}I-labelled toxin, they have been able to follow toxin entry. Trypsin treatment was used to distinguish surface-bound toxin (accessible to proteolytic enzyme) from internalized ligand (protected from proteolysis and recovered with fibroblast cell pellet). Both internalization and subsequent intracellular trafficking were followed at the ultra-structural level by the biotin-avidin system. In this system biotinyl toxin is bound to cells at 4°C for 30 min after which avidin-stabilized gold colloids are added at 4°C for 30 min. The cells are washed and warmed for an appropriate length of time, allowing the internalization of toxin-gold complexes, and then processed for electron microscopy (EM). Both the trypsin release assay and the EM approach showed that one half of the PE prebound to mouse fibroblasts was internalized by 5 minutes after warming to 37°C. The PE was routed in endosomes to the Golgi within 5-15 minutes after entry, and was transported to lysosomes by 30 minutes. In brief, after binding to surface receptors, the toxin rapidly clusters into specialized organelles of the plasma membrane called clathrin-coated pits, and is then internalized into the endosomes, passes through the Golgi region, and ultimately is delivered to the lysosomes (see figure).

The acidic environment of the endosome seems to be important in the expression of toxin activity. If the pH of endosomes is raised the toxic activity is decreased. Such basic agents as methylamine and ammonium chloride have been shown to alter the normal routing of ligands internalized by receptor-mediated endocytosis (RME). Ultrastructural evaluation of the trafficking of PE gold, in the presence of either of these two agents, revealed that entry via clathrin-coated pits was blocked and that PE gold was sequestered in large electron-translucent vesicles. However, the toxin still could be internalized. This implies that in the presence of basic amines, receptor-bound PE is entering by an alternate but ineffective route.

In summary, Dr. Saelinger's data suggests that the sensitivity of cells to PE is the result of the specific interaction of toxin with its cell surface-binding proteins. After the initial interaction, the toxin is internalized by receptor-mediated endocytosis and is routed to the Golgi via endosomes. Toxin internalized by alternate routes, i.e., nonspecific adsorptive pinocytosis, is nontoxic because it is delivered directly to the lysosome where it is rapidly degraded. Based on this intracellular routing of toxin, Dr. Saelinger feels that she has described the requirements for the construction of an ideal chimeric toxin. Chimeric toxins, a term synonymous with hybrid toxins, immunotoxins and magic bullets, are heteroconjugates

formed between various protein ligands and toxic proteins or fragments derived therefrom. These fascinating studies could lead to the development of more effective therapies for Pseudomonas infections.

Yvonne T. Maddox

Yvonne T. Maddox, Ph.D.

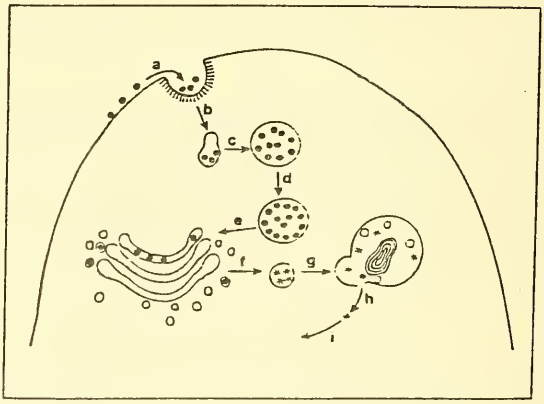


Fig. 1. Internalization of PE (●) by mouse LM fibroblasts. a: PE enters via RME through clathrin-coated pits (30 s); b: PE transferred to endosome (1 min); c: endosomal coalescence (5 min); d: continued endosomal coalescence (10 min); e: transit to the Golgi (15 min); f: putative transit from Golgi to endosome with conversion to enzymatically active form (★); g: fusion of putative post-Golgi endosome with lysosome (20 min); h: escape of active PE by leaky fusion (20 min); i: NAD-dependent ADP ribosylation of EF-2.

Memorandum

Date January 11, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject Research Highlight: Bacterial Translocation of the Gut Flora, GM36376 (Deitch, Edwin; LSU Medical Center, Shreveport)

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program, NIGMS

Bacterial infection is a frequent and serious problem in patients who survive the initial shock phase of thermal injury. Infections in burn patients generally have been considered to arise from exogenous organisms colonizing the burn wound. Consequently, in considering the therapy of burn patients the emphasis has been on the adaptation of strict infection control policies to reduce contact of the patients with exogenous organisms, and the use of topical antimicrobial agents to reduce bacterial colonization of the wound site. Until recently, very little attention has been directed to the indigenous microflora in the gastrointestinal (GI) tract of the burn patient as a source of the repeated episodes of bacteremia; yet the gut may serve as a reservoir for organisms causing bacteremia in the burn patient. Support for this hypothesis is that the presence of a gut flora appears to increase mortality after a thermal injury, since germfree mice survive a thermal injury better than conventional mice. Additionally, when the GI tracts of germfree mice are colonized prior to a thermal injury with bacteria such as E. coli, bacteremia occurs and the burn-related mortality of these animals is increased. It has been speculated that the GI tract may cause the infections; either indirectly by fecal organisms contaminating and colonizing the wound, or directly by passage through the intestinal wall.

Dr. Deitch and his co-investigator, Dr. Rodney Berg, have been investigating the relationship between the GI microflora, systemic host defenses, and trauma in an attempt to delineate the mechanisms by which bacteria contained within the GI tract can cross (translocate) the GI epithelium to cause systemic infections. They use the term "bacterial translocation" to describe the passage of viable indigenous bacteria from the GI tract through the epithelial mucosa to the mesenteric lymph nodes (MLN) and other systemic organs. These investigators have identified several factors that promote bacterial translocation: overgrowth of certain bacteria populating the GI tract, deficiency of the host's immune defense mechanisms, and physical disruptions of the GI tract mucosal barrier. A thermal injury, either directly or indirectly, may cause a disruption or impairment of any of these mechanisms.

In early studies, using a rat model, Dr. Deitch sought to determine whether nonlethal thermal injuries of increasing severity would promote the

translocation of indigenous bacteria from the GI tract (In general, Gram-negative enteric bacilli, Gram-positive cocci, and lactobacilli make up the indigenous bacterial flora of the GI tract.) The rats received 20% to 40% body surface area (BSA) scald burns and were sacrificed at 1, 2 and 21 days post burn, and various organs were cultured quantitatively for aerobic, anaerobic, and facultative anaerobic bacteria. No viable indigenous bacteria were cultured from the organs (spleen, liver, blood and MLN) of control rats or from rats receiving 20% BSA burns. However, 44% of the rats receiving 40% BSA burns exhibited viable bacteria in their MLN 2 days after thermal injury. Since bacterial infections of burn patients are often caused by Gram-negative enteric bacteria, the investigators tested bacterial translocation of E. coli to other organs in rats with 40% burns. E. coli translocated at a greater incidence than other bacteria to the spleens, livers, and peritoneal cavities of the burned rats. By 14 days after thermal injury, the rats were able to clear the E. coli from their spleens and livers, but the translocated bacteria remained localized in the MLN. These results support the concept that the indigenous GI flora or exogenous organisms colonizing the GI tract, are potential sources of septicemia after thermal injury.

Dr. Deitch, in looking for potential explanations for the burn-induced increases in bacterial translocation, has recently focused on the effects of endotoxin on translocation. Endotoxin has a wide range of biologic actions including the ability to depress host antibacterial defense systems, increase vascular permeability, impair cellular metabolism and oxygen utilization and produce profound hemodynamic changes, resulting in hypotension and death. The investigators chose to study endotoxin effects, because prior studies suggested that endotoxin increased the permeability of the gut mucosa to bacteria, and impaired host immune defenses. Furthermore, endotoxemia is common after thermal injury and is associated with conditions leading to multiple organ failure.

The results of these experiments documented, for the first time, that nonlethal doses of endotoxin promoted bacterial translocation to the MLN in a nonlethal burn injury (25%) and unburned mice, but only in burned mice did the bacteria translocate from the gut to the other systemic organs. Furthermore, the mortality rate of mice receiving only endotoxin or burns was less than 10% while the combination of endotoxin plus thermal injury increased the mortality rate to 100%. These results indicate that endotoxin will promote bacterial translocation from the gut of unburned healthy mice with a normal indigenous GI microflora; however, the translocating bacteria are limited to the MLN. In contrast, burned mice receiving endotoxin developed systemic bacterial invasion of multiple organs and ultimately died of uncontrolled systemic sepsis.

The mechanisms whereby endotoxin promotes bacterial translocation with septicemia and death are not known. However, Dr. Deitch's results suggest that endotoxemia, whether derived from the burn wound, GI tract or visceral sites of infection, such as the lung, may initiate a cycle of endotoxin-mediated bacterial translocation, which may be a source of lethal infections after thermal injury. Dr. Deitch is currently conducting studies

to determine which components of the host defense system (neutrophils, immunomodulators, etc.) are important in preventing translocation as a prelude to developing therapeutic strategies for preventing systemic infections by organisms which colonize the GI tract of burn patients.

Yvonne T. Maddox, Ph.D.

Memorandum

Date September 28, 1989

From Program Administrator, Biophysics and Physiological Sciences Program, NIGMS

Subject RESEARCH HIGHLIGHT: Cellular and Molecular Pathobiology of Wound Repair, 5 R01 GM37223-03 (Grotendorst, G.), Univ. of South Florida

To Director, NIGMS
Through: Director, BPS Program, NIGMS *Jacque*

Tissue regeneration and repair appear to be controlled in part by specific peptide factors which regulate the migration and proliferation of the cells involved in the repair process. Some of these factors have been isolated and identified, such as platelet derived growth factor (PDGF), transforming growth factor-alpha and-beta (TGFα, TGFß), epidermal growth factor (EGF), and fibroblast growth factor (FGF). There are a variety of cellular sources for these factors, including the cells which are an integral part of the cellular machinery of wound repair. For example, platelets release PDGF, TGFß and an EGF during aggregation. Activated monocytes (macrophages) release PDGF-like activity and express the c-sis gene which codes for the PDGF-ß chain and produce FGF. In addition, macrophages also secrete TGFß and express the TGFß gene. Endothelial cells and smooth muscle cells both express the c-sis gene and secrete a PDGF-like activity as well.

Dr. Grotendorst has been focusing his attention on PDGF and TGFß and their role in wound healing, and in preliminary studies, he has used purified growth factors in an in vivo wound model. To study the role of growth factors during the repair process, Dr. Grotendorst uses stainless steel mesh chambers as part of a model wound system developed by Dr. Thomas Hunt of the University of California, San Francisco in 1967. These chambers, when implanted subcutaneously, induce a tissue repair response which initially fills with fluid (wound fluid). The specific cell types which function during the repair process subsequently invade the fluid in the chamber and release various protein mediators. Hence, it is possible with this model to follow the amounts of various chemotactic and/or mitogenic factors at the wound site.

Dr. Grotendorst sought initially to compare the levels of PDGF and TGFß in wound fluid collected from normal and healing impaired animals. In these initial studies, he observed that an intravenous injection of an anti-tumor drug, adriamycin, prior to the implantation of the wound chambers induced a significant reduction in new connective tissue formation whereas the same

dosage given after the repair process was underway had no effect on the amount of connective tissue formed within the chamber. Dr. Grotendorst has suggested that the adriamycin model is an impaired wound healing model and the results certainly suggest this to be the case. Specifically, the data showed that animals which received a single injection of adriamycin 4 days prior to wounding exhibited an 80% reduction in the amount of new collagen deposited 21 days after implantation of the wound chamber. In marked contrast, animals which received the same dosage of adriamycin 7 days post implantation exhibited no change in the amount of connective tissue deposited by day 21 and was comparable to the non-treated control group. These data indicated that the adriamycin is acting indirectly to inhibit the repair response. There are some known toxic effects of adriamycin which may help explain the mechanisms by which wound repair is inhibited. One of the tissues which is most sensitive to adriamycin is the bone marrow, and it is well known that patients receiving adriamycin therapy as an anti-tumor drug become thrombocytopenic and leukopenic. When the blood of the rats used in Dr. Grotendorst's studies were analyzed for levels of circulating platelets and leukocytes, these cell counts were significantly lowered after the animals received a single injection of adriamycin. Thus, the adriamycin may be affecting the repair process by abrogating the inflammatory phase of the response, as the number of platelets and leukocytes present are not sufficient to induce a repair reaction.

Platelets and macrophages seem to be the primary candidates as the sources of PDGF and other growth factors during tissue repair; therefore, Dr. Grotendorst compared the levels of both PDGF and TGFß in the wound fluid collected from chambers which had been implanted into normal animals with that present in chambers which had been implanted into the healing-impaired animals (day-4 adriamycin treated). Using the smooth muscle cell chemotactic assay which is specific for PDGF and the anchorage independent growth of natural killer cells assay for TGFß activity, the levels of these growth factors were determined. Interestingly, the concentrations of both PDGF and TGFß varied during the course of the repair response, with the peak levels present at the same time as the greatest increase in connective tissue deposition in the wounds (day 21). The amount of these factors present in the fluid collected from adriamycin treated (day-4) animals was reduced from that seen in the normal animals. Dr. Grotendorst has now completed studies which show that supplementation of the wound chambers in the healing impaired animals with pure PDGF and TGFß restores the amount of new collagen deposited to that seen in normal animals.

Dr. Grotendorst and his colleagues are now using specific cDNA probes for the PDGF and TGFß genes to determine whether the genes coding for these proteins are expressed at the wound site. The

current findings suggest that these growth factors may serve as initiators which induce the formation of new connective tissue. During normal repair processes these factors would be released by platelets, macrophages, or possibly other cell types, and would recruit the connective tissue cells from the undamaged surrounding tissue to the site of injury. There is indeed much optimism and some speculation as to whether these factors might have a practical use in clinical situations of compromised wound repair and in understanding fibrotic diseases in which uncontrolled connective tissue proliferation occurs. Wound healing is a very exciting area, and Dr. Grotendorst's molecular and cellular studies serve as an excellent base for future investigations directed at understanding defects in wound healing.

Yvonne T. Maddox, Ph.D.

Memorandum

Date September 27, 1989

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Synthesis and Testing of Living Skin Equivalent, GM28485 (Stevenson, T.), Univ. of Michigan, Ann Arbor

To Director, NIGMS

Through: Director, BPS Program, NIGMS

The successful conversion of an open wound to a closed one remains the major goal in the treatment of large burns. Any delay in this conversion permits continued fluid loss, prolongs healing, and increases the likelihood of infection. In patients with massive burns, donor sites from the patient's own body surface are often exhausted before all the burn wounds can be covered. Subsequent healing of the donor sites will allow these sites to be used for grafting again in about two weeks. Unfortunately, the number of times a site may be used is limited.

This limitation has prompted investigations into the use of cultured autologous (from same organism) keratinocytes for wound closure. Essentially, the process involves isolating keratinocytes from the epidermal layer of the skin, growing them in culture and transferring the cultured cell sheet to the wound site. Enzymatic digestion is required to separate the fragile keratinocyte sheet from the tissue culture substrate. Problems occur because this sheet is subject to disruption and must be handled carefully, making it extremely difficult to transfer each sheet to the burned area.

Dr. Stevenson has conducted studies to improve cell growth in culture and to improve the transfer of keratinocyte sheets. He began these experiments by using keratinocytes taken from neonatal female rats. The keratinocytes were isolated and suspended in either a control medium containing normal Ca^{++} (1.2mM) or in a supplemental medium consisting of known growth accelerators and low Ca^{++} (0.035mM). These two groups of keratinocytes in culture were plated on a 100μm thick sheet of Type I bovine dermal collagen. Dr. Stevenson has referred to the combination of the collagen sheets and the viable keratinocytes as quasi-skin. The cells were grown in culture under the usual conditions (i.e., 37°C, 5% CO_2 with humidity; provided standard nutrients 3 times per week).

The proliferation rate of the cells was assessed by measuring the rate of incorporation of ^{3}H-thymidine into the replicating DNA on

days 3, 7, 10, and 20 after plating. To ascertain the ability of the keratinocytes to survive grafting, Dr. Stevenson used the quasi-skin from each of the two culture media on days 7 and 14 to cover full thickness wounds/defects (30mm in diameter) created on the dorsum of adult male rats. These animals were examined routinely, and postoperative biopsies were taken of the wounded areas and of the tissue immediately adjacent to the defect site.

The results obtained from these experiments are extremely interesting. For cells grown in the control medium, the rate of ^{3}H-thymidine incorporation was high on days 3 and 20 and low on days 7 and 10. For the cells grown in the supplemental medium, the levels were consistently high at each time point. Dr. Stevenson interprets these results as indicative of a role for low-calcium in maintaining the proliferation rate of the keratinocytes. He suggests that under conditions of normal calcium, incorporation of ^{3}H-thymidine may be depressed on days 7 and 10 due to stratification and differentiation of the keratinocytes.

The light microscopy showed the wound defects to be completely re-epithelialized by day 45. Quasi-skin grown in either of the growth medium appeared to be as effective for re-epithelialization. Barr bodies were observed in some of the nuclei, representing the female cells growing on the male recipient and also indicated the persistence of the transferred cells. There was no gross difference in the rate of wound healing using cells grown in either culture medium, and the length of time in culture did not seem to affect wound healing.

An important result from these studies, and one with enormous clinical potential, was the creation of a stable collagen sheet with attached viable keratinocytes. Indeed, the use of the collagen sheet as a growth and carrier substrate produced a compound of increased mechanical stability compared to a sheet of keratinocytes alone. The quasi-skin was stable enough to lift with forceps for transfer to the wound surface.

There were no visible remains of the collagen sheet when the tissue was examined under the microscope at either 45 days or 60 days. This suggests that, while the collagen sheet was one that was cross-linked enough to withstand the collagenase which is produced by the keratinocytes in culture, it was not cross-linked enough to withstand long-term exposure to the wound environment. It is important for collagen to remain long enough to anchor fibrils and direct fibroblast ingrowth so that repair is complete. Dr. Stevenson is planning future studies to cross-link collagen to withstand longer term degradation.

There is also the question of whether the 30mm diameter wounds employed in this rat model will really predict usefulness in

larger wounds where adherence and wrinkling of the collagen sheet may become substantial problems. Also, further studies need to be done to determine the optimum duration of keratinocyte growth in culture before transfer to the excised wound. Nevertheless, the technique of using a collagen sheet seeded with keratinocytes may hold good possibilities, and these observations must now be transferred to the study of human keratinocytes grown in culture to determine therapeutic importance.

Yvonne T. Maddox

Yvonne T. Maddox, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 29, 1988

From Program Administrator, Biophysics and Physiological Sciences (BPS) Program, NIGMS

Subject RESEARCH HIGHLIGHT: Cryopreservation of Organs, 2 R01 GM17959-19 (Meryman, H.), American Red Cross

To Director, NIGMS

Through: Director, BPS Program, NIGMS
Chief, Physiological Sciences Section, BPS Program, NIGMS

Cryopreservation is an extremely important tool for modern medical science; a variety of single cells (blood, sperm, ova, bone marrow) and simple tissues (cornea, skin, embryos) can now be stored for long periods of time at deep subzero temperatures (generally -196° C in liquid nitrogen) with successful resumption of function after thawing and transplantation. The growing use of organ transplant therapy as a critical intervention strategy fuels the drive to extend freezing preservation to larger intact organs (i.e. kidneys). Presently, the technique of organ transplantation relies upon the use of hypothermia (packing in ice) to provide a few hours of viability for the transfer between donor and recipient.

The field of cryopreservation has not made rapid advances. In fact, the modern era of cryopreservation began in 1949 with the work of Polge which demonstrated the successful revival of spermatozoa after freezing in the presence of high concentrations of glycerol as a cryoprotectant. Since then the field has largely advanced by empirical experimentation, manipulating cooling and warming rates or types and amounts of cryoprotectants, in an attempt to optimally compensate for the various types of freezing injuries that have been identified. The most damaging of these is ice formation which has damaging consequences including structural damage of extracellular components, osmotic shock due to the rapid redistribution of water and solutes across the cell membrane during freezing and thawing, damage to membrane and/or subcellular structure when a critical minimal cell volume is exceeded, damage to cell structure or function due to dehydration and elevated solute/ion levels in remaining liquid compartments, and metabolic damage due to the ischemia and anoxia imposed by the frozen state. In addition, the agents added as cryoprotectants often have pronounced toxic effects on cellular metabolism.

Dr. Harold Meryman has been studying the use of vitrification in lieu of freezing as an alternative approach to organ cryopreservation. Vitrification is the solidification of a liquid brought about, not by crystallization, but by an extreme elevation in viscosity during cooling. During vitrification the solution is said to become a glass and all molecular motions are halted. Vitrification of relevant aqueous solutions using cooling rates that are realistic for whole organs requires the presence of high concentrations of

a cryoprotective agent. The primary challenge that must be met in order to successfully vitrify organs is to make the required concentrations of cryoprotectant nontoxic to the organ.

The basic concepts of vitrification are best described by reference to the attached figure which represents a hypothetical cryoprotectant. T_m is the equilibrium freezing or melting point curve. Solutions normally supercool to some point between T_m and T_h (the homogeneous nucleation temperature) before they actually nucleate or begin to freeze, as represented here by X's. T_g is the glass transition temperature, at which a supercooled liquid vitrifies. T_d is the devitrification curve, at which the previously vitrified solution freezes upon rewarming.

Several solutions have been used as cryoprotective agents and include ethylene glycol, glycerol, and dimethyl sulfoxide in varying concentrations. Cryoprotectant toxicity and osmotic injury are influenced by the concentration of cryoprotectant and the time and temperature of exposure.

One method used to reduce the effective amount of cryoprotectant required for vitrification is to apply high hydrostatic pressure. High pressures lower T_h and elevate T_g, thus shifting the point of intersection to a lower concentration of cryoprotectant. There are certain rather well-defined regions on the phase diagrams in which different types of vitrification behavior appear.

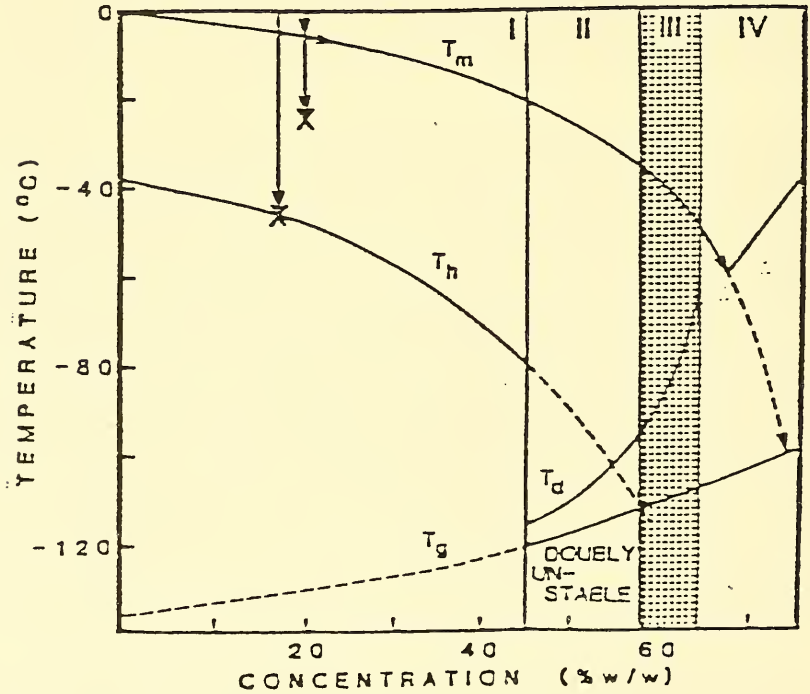
In Region I, which is a relatively dilute area, vitrification is almost impossible because heterogeneous and homogeneous nucleation are unavoidable. In Region II, which is the more concentrated region, both types of nucleation are inhibited, as is crystal growth. Although it is possible for vitrification to occur in Region II, glasses formed here are thought to be heavily nucleated and normally are fated to freeze or devitrify upon warming. At still higher concentrations (Region III), T_h becomes equal to and then actually falls below T_g. In this region it is possible to slowly cool liquids directly to T_g without experiencing any detectable freezing events, despite the presence of heterogeneous nucleating agents. The intersection between the T_h curve and the T_g curve therefore established the threshold or lowest possible concentration of cryoprotectant that might be used for organ vitrification. Although it is possible to vitrify organs in this region without forming detectable quantities of ice, the existence of devitrification (ice crystallization during warming) upon warming has been interpreted as evidence for significant heterogeneous nucleation during cooling. Finally, at Region IV, the devitrification curve vanishes even at slow warming rates. Here all nucleation is prevented, and the system is virtually stable. Although ideal in principle for organ preservation, this region appears to be well beyond reach due to overwhelming problems with cryoprotectant toxicity. Therefore, Region III is the main focus of interest for preservation. Once the temperature has been brought to below T_g, the pressure can be released without danger of crystallization.

As mentioned earlier, devitrification has been a major obstacle and this phenomenon and its relationship to the composition of vitrification solutions, to pressure, time and to temperature is the current lead being investigated by Dr. Meryman. Hydrostatic pressure inhibits the freezing of water and can

lower the concentrations of cryoprotectants needed for vitrification. Small specimens have been successfully vitrified and rewarmed, but scaling the procedure up to handle large organs requires electromagnetic heating at high energy levels.

Other than cryoprotectant toxicity, devitrification is the single most imposing primary barrier associated with organ cryopreservation. Dr. Meryman has determined that a warming rate of about 1000°/min will suppress devitrification. He has also demonstrated that ethylene glycol is the easiest cryoprotectant to add to and remove from tissues. He is currently focusing this technique on the preservation of rabbit kidneys and measuring organ function following acute or permanent transplantation studies. The ultimate aim of this research is to achieve the cryopreservation of human organs and vascularized tissue.

Yvonne T. Maddox, Ph.D.



Phase diagram of a Hypothetical Cryoprotectant (Fahy, G., MacFarlane, D., Angell, C. and Meryman, H., Cryobiology 21:407-426, 1984).

CELLULAR AND MOLECULAR BASIS OF DISEASE PROGRAM

The Cellular and Molecular Basis of Disease (CMBD) Program continues to expand in several directions. Although a significant part of the program is devoted to biochemistry, including such areas as enzyme reaction mechanisms and control of complex metabolic pathways, a larger part of CMBD can be subsumed under the heading "Cell Biology." Although historically "cell biology" described a relatively narrow set of research endeavors, many of which were not very quantitative, that is no longer true. In a sense a large fraction of basic biomedical scientists are asking questions at a cellular or organellar level. We want to know what the role of specific biochemical processes is in the complex living cell. Modern researchers freely borrow technology from the entire range of biomedicine to answer these questions. It is difficult to find a lab that is not employing modern molecular biological techniques, particularly recombinant DNA technology, to address such cell biological problems. Additionally, the major advances in structural biological techniques, taking the field out of the exclusive domain of its traditional practitioners, have encouraged researchers to ask questions about molecular structure and the relationship between such structure and physiological function. The result has been an explosion in the number of grants coming to CMBD. As one peruses the highlights that follow (which are arranged roughly from the most molecular to those asking questions about complex cellular behavior and cell-cell interactions), the results of this explosion become evident. Advances in physical measurements coupled with recombinant DNA manipulations have created expanded opportunities in the analyses of enzyme mechanisms. A broad multidisciplinary attack, including genetic, biochemical, molecular biological, and virological approaches, combined with an exceptionally talented cadre of scientists, has made major inroads in our understanding of protein sorting and trafficking. A set of seminal findings, using genetic and biochemical strategies, has provided major breakthroughs in our understanding of cell cycle control. Finally, advances on this broad front have revealed new insights into disease processes and, as is evident from several highlights provided, suggested or even specified therapeutic strategies.

The highlights focusing on enzyme mechanisms demonstrate the use of multiple strategies to further our understanding of enzyme mechanisms, as well as the far-ranging importance of our broadened understanding. For example, the crucial technology in the work of Dr. Jeremy Knowles is selective alteration of specific amino acids using site-directed mutagenesis. Clever combination of such mutagenesis with modern improvements in X-ray crystallography have allowed Dr. Evan Kantrowitz to analyze complex enzymes. Dr. James Paulson has benefited from the broad progress in protein and enzyme purification to allow him to explore new members of an important class of enzymes. The fruits of these efforts have been evident. For example, Dr. Knowles has succeeded in changing the specificity of a given enzyme, thus opening up the possibility for designer catalysts to be used in industrial or therapeutic processes. Dr. Christopher Walsh has pioneered the exploration of a unique class of enzymes not found in animals, namely those requiring nickel for catalysis. These enzymes, although few in number, are of enormous importance in the "metabolism" of the biosphere. The isolation and purification of

sialyltransferases from cells by Dr. Paulson will now lead the way to studies of the role of tissue-specific sialic acids in essential cell processes, as detailed in the "Bullet" provided. Finally, some of the most interesting biochemistry has concerned some simpler molecules. Dr. Bennett Shapiro has discovered a new family of low-molecular-weight antioxidant compounds he has named "ovothiols." They are exceptionally potent reducing agents which have been shown to play a major protecting role in the cells of some lower organisms and may also play such a role in mammalian cells.

Again, for this report, we have described continuing advances in the study of cellular trafficking, particularly of proteins. The basic questions remain: how do cells synthesize these special membrane or secretory proteins? how do proteins travel to specific parts of the cell? how do soluble proteins cross membranes, for example in mitochondria? One of the more intriguing sets of studies is that of Dr. Martin Snider concerning the cycling of surface glycoproteins and their receptors. He and others have shown that the mannose-6-phosphate receptor serves to target lysosomal enzymes both into the lysosome from the Golgi and into targeted cells after release into the circulating blood from other distant sites. Another interesting finding in the area of protein trafficking is the elucidation of the DNA sequence of a gene, sec 4, crucial to secretory processing in yeast. This gene turns out to have extensive homologies with the ras gene, a yeast version of a mammalian oncogene. This generates some intriguing questions about the possible similarities in the secretory pathway and that of neoplastic transformation. Some of the most singular contributions have been those of Dr. Peter Walter. Essentially, Dr. Walter has extensively characterized the biochemical steps and essential components required to synthesize membrane-bound and secretory proteins on rough endoplasmic reticulum. Lastly, some recent work by Dr. Ari Helenius and colleagues, not reported in a highlight, has elucidated some surprising properties of the initial steps in protein processing. Essentially what they have shown is that the cell imposes strict quality control on the molecules it transports. The initial folding of the subunits and assembly into oligomers are both crucial for transport competence. Specifically, wild-type hemagglutinin can be transported and subsequently secreted. Any amino acid substitutions that change the molecular folding of either the soluble or the membrane-bound hemagglutinin form aggregates that stay in the endoplasmic reticulum.

The cyclin story is one of the most exciting to come out of cell biology in the decade. The question of the control of cell division and the cell cycle is a longstanding one and has been the subject of decades of investigations with modest insights into the basic mechanisms involved in cycle control. The question is an important one; after all, such cellular control is essential in embryogenis, development, and aging, as well as in disease processes such as tumorigenesis and metastasis. The highlight entitled "Overview of Cell Cycle Control" provides just that, a useful overview. Briefly, cell cycling appears to be controlled by the level of a specific protein complex, MPF. This complex is composed of two critical components, cyclin and a specific kinase. The interactions of these components, as well as their interactions with substrates or as substrates, are the focus of major efforts in laboratories throughout the world since their recent initial discovery.

There are, of course, many other foci of effort which have generated excitement throughout the biomedical community. We will single out three. First, as revealed in the highlights about the works of Drs. Robert Simoni, Tom Stevens, Ching Kung and others as well as the Council's approval of MERIT awards to Drs. Carolyn Slayman and Simoni, molecular dissection of membrane transport has proven very successful. Site-specific alteration of transport subunits combined with specific inhibition and spectroscopic measurements have provided us with critical information about the active sites and transport mechanisms of membrane transporters. Secondly, the exciting discovery of kinesin, a new intracellular force-generating transport molecule, has been followed by a series of related discoveries, as evidenced by the highlights and bullets on the work of Drs. James Spudich and Richard Vallee. We are discovering more and more classes and varieties of cellular ATPase involved in intracellular transport. Lastly, we want to emphasize some of the major progress in the analysis of the structure and function of gap junctions. The laboratories of Drs. Daniel Goodenough and David Paul at Harvard have been mapping the structure of the gap protein, connexin, and elucidating the specific domains (intracellular, extracellular, and intramembranous) of the molecule. These labs also have recently cloned a connexin gene from frog oocytes and showed that it has a unique role in early development.

As in years past, CMBD grantees are discovering new approaches to therapy and prevention while pursuing the most basic research. Although in general there is a lag time between basic findings and medical applications, there frequently are specific serendipitous discoveries that lead directly to some immediate medical application or provide a surprising insight into some disease state. For example, investigators working on serine proteases and its inhibitors (see highlights about Drs. Dennis Cunningham and Huntington Potter) have revealed a close link between deposits of serine protease inhibitor and Alzheimer's disease. This has led to formulation of one of the leading hypotheses of the etiology of Alzheimer's. Dr. Dennis Carson, as described in another highlight, studied a genetic defect resulting in a serious immunodeficiency. This result appears to be leading to an exciting new chemotherapy for treatment of chronic lymphocytic leukemia. In another example, work we support on the understanding of proton-driven ATPase, combined with studies on basic endocytotic processes and lysosomal biochemistry, have served as the basis for important findings on the metabolism and biology of the pernicious parasite Leishmania (see highlight). The investigation by Dr. Martin Snider into the role of mannose-6-phosphate receptors has already been used to develop a new enzyme by a biotechnology firm for treatment of the serious genetic disorder, Gaucher's disease. Recent work by Dr. Eugene Butcher (see highlight) has resulted in the identification of specific receptors on endothelial cells (which line blood vessels) for attracting lymphocytes. This system appears to play a role in normal and pathological conditions, such as inflammation, and autoimmune destruction of tissues. Finally, in addition to the highlights included in this report, grantees continue to make discoveries of medical interest. Recently, Dr. Alan Finkelstein and colleagues, mainly at Albert Einstein Medical School, have discovered that anthrax toxin can form channels in membranes. These channels may be the key to the mechanism of entry of the toxic portions of the toxin (such as EF, which has been shown to be a potent calmodulin-dependent adenylate cyclase) into the cytosol. This property of the entire toxin appears to be analogous to the mechanisms of entry of diphtheria, tetanus, and

botulism toxins. It also opens up a possible strategy for prevention of the entry of the toxin into the cell.

For this year's biennial report of CMBD activities we provide an informal compendium of highlights with a brief accompanying commentary. The highlights serve to outline the range of significant research we support. They are just a representative cross-section of the exciting science going on nationally in what is considered the age of biomedical research.

The organization of CMBD is unchanged in the past 2 years, and the only change in professional staff has been the loss of Dr. Lore' Anne McNicol. There have been no basic policy changes in research or training. Our fiscal fortunes in both research and training have been somewhat rocky, nevertheless, we have continued to fund many excellent research and training grants.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7. 1988

From Program Administrator, CMBD

Subject Research Bullet: Mechanisms of Enzyme Catalysis
GM 21659-15 (Jeremy R. Knowles), Harvard University

To Director, NIGMS
Through: Director, CMBD

Investigators are combining our considerable knowledge of enzyme structure and function with powerful new molecular biological techniques to produce enzymes having different catalytic properties. These efforts have at least two goals in mind. One goal is to provide a new enzyme with a different substrate specificity. This goal presents obvious practical possibilities, in the laboratory or industrial production of new chemicals or of chemicals that otherwise can be produced only with great difficulty. A second goal for producing enzymes with altered specificity is to learn more about the structural components in enzymes that enable them to have such great catalytic power. This latter goal is the driving impetus for the research efforts of Dr. Jeremy Knowles (GM 21659) of Harvard University.

For several years, investigators have used mutagenesis coupled with strong selection procedures to generate enzymes with altered catalytic specificity. For example, in 1975, Dr. Knowles and his co-workers were able to use mutagenesis and selection procedures to generate an enzyme that destroyed cephalosporin, instead of penicillin as did the original enzyme. Other investigators have used procedures like these to change a dehydrogenase's preferred substrate from ribitol to xylitol. Still another interesting case, in this instance a selection that occurred through natural processes, was the appearance of an enzyme that could hydrolyze nylon, an instance that occurred through natural processes in the waste water from nylon plants. Investigators also have been able to increase the catalytic activity of some forms of isomerase.

In the past, researchers attempting to use these observations to understand the structure and function of enzymes have been hampered by a lack of structural information about the changes in enzyme structure that led to altered catalytic or binding properties produced by mutagenesis. Dr. Knowles is now using a combination of sophisticated molecular biological, kinetic, and structural sophisticated techniques to attempt to correlate structural with functional changes in several important or well-known enzymes.

These studies not only will provide new insights as to how enzymes structures originated and how they cause such a great enhancement in chemical reaction rates, but they may also someday provide the groundwork for producing new industrially useful enzymes that generate molecules which either do not exist today or that can be produced only with great difficulty.

M. Janet Newburgh

M. Janet Newburgh, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Program Administrator, CMBD

Subject Research Bullet:Nickel-Dependent Enzymes in Methanogenesis
GM 31574-05 (Christopher T. Walsh), Harvard Medical School

To Director, NIGMS
Through: Director, CMBD

Methanogenic bacteria utilize carbon dioxide and hydrogen gas to produce methane, otherwise known as natural gas. These bacteria are members of the Archaebacteria kingdom and play a crucial role in the anaerobic digestion of the biomass. (Every year, one billion tons of methane are produced in the biosphere, and essentially all of this is due to methanogenic bacteria.) Dr. Christopher Walsh (GM 31574) and others have been studying the structures and mechanisms of the enzymes in methanogenic bacteria that catalyze the production of methane. This research has shown that these bacteria contain three enzymes which require nickel ions for catalytic activity. Although urease, a plant enzyme, was already known to contain nickel, the three bacterial enzymes are the first examples of a direct functional role of nickel in biological oxidation-reduction reactions.

A common feature in the nickel-containing oxidative-reductive enzymes is the presence of other cofactors that probably are important in catalysis.

The nickel-dependent hydrogenases, which reduce a substrate such as sulfur by obtaining electrons from hydrogen gas, are now known to contain iron-sulfur clusters and flavin adenine nucleotide, in addition to the nickel. Methyl coenzyme M Reductase, which catalyzes the actual release of methane gas from a sulfur-containing complex, has been found to contain a nickel-containing tetrapyrolic factor, known as F-430. Nickel-containing carbon monoxide dehydrogenase, which catalyzes the interconversion of carbon monoxide with carbon dioxide, has been much more difficult to study. It probably also contains iron-sulfur complexes, possibly in close association with the nickel. Clearly, much remains to be learned about the nickel in these enzymes, including the changes in electronic structure that occur during catalysis.

In addition to the basic scientific interest of research on the nickel-containing enzymes, current thinking is that the enzyme catalytic centers may resemble nickel-based catalytic reactions which are used in laboratory and industrial settings. Thus, further study of the structures of the catalytic centers of these enzymes and the electrochemical role of nickel during enzyme catalysis may also provide the necessary groundwork for the design of better organometallic catalysts.

M. Janet Newburgh

M. Janet Newburgh, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 8, 1988

From Program Administrator, BPS Program, NIGMS

Subject Research Bullet: 2 R01 GM27904-09 (Paulson, James C.
Mammalian Sialytransferases - UCLA)

To Director, NIGMS
Through: Program Director, CMBD

Cell surface sialic acids are known to mediate the adhesion of a variety of toxins, plant and animal lectins, antibodies and viruses. In addition, sialic acids have been implicated as important determinants in the cellular receptors for killer cells, interferon, and recirculating lymphocytes. Sialic acids also have important masking functions. For instance, desialylated serum glycoproteins are rapidly cleared from the blood by carbohydrate specific lectins found in the liver and there is a strong correlation between between the degree of sialylation of melanoma cell lines and their propensity for metastasis. The structural diversity of the sialosides together with the terminal positions of the sialic acids make them ideal structures to participate in biological recognition.

At least 12 different sialyltransferases are required to synthesize all the sialic acid containing carbohydrate sequences produced by mammalian cells. To date only four sialyltransferases have been purified to homogeneity. Just recently, the first sialyltransferase was cloned and completely sequenced by Dr. James Paulson at UCLA. In fact, this is the first glycosyltransferase ever to be completely sequenced. The same or analogous procedures are likely to be successful in isolating and sequencing the other sialyltransferases. With these transferases available in quantity and of known structure, it will be possible to obtain new information on the separate regulation of the individual transferases and their role in biological recognition. It should bring us a step closer to understanding the organization of the cellular glycoyslation machinery.

Lee Van Lenten, M.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

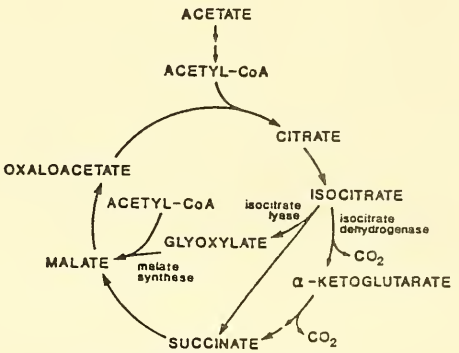
Date September 12, 1988

From Chief, Molecular Section, CMBD, NIGMS

Subject A Research Highlight: "Regulating the Use of the Glyoxalate Bypass." Dr. David LaPorte (GM 33927) University of Minnesota

To Director, NIGMS
Through: Director, CMBD

E. coli, like most organisms, utilizes the well-defined Krebs cycle to derive energy from acetyl-CoA, the product that results in common from the catabolism of many fats, carbohydrates, and other substances. On net, both carbons of the acetate that enters the Krebs cycle are lost as CO_2, and are thus unavailable for biosynthetic purposes. However, E. coli will grow on acetate (another metabolic precursor of acetyl-CoA) as the sole source of carbon. Growth, of course, requires the biosynthesis of many carbon-containing materials, so it is clear that the cell has the capacity to avoid the loss of carbon that is inherit in the Krebs cycle. This capacity has its basis in the glyoxalate bypass, which, as the figure below shows, provides a metabolic pathway around the two CO_2-generating steps of the Krebs cycle. The use of this bypass means that acetate need not be lost as CO_2 but rather can be



Krebs cycle and glyoxylate bypass

employed in the net production of oxaloacetate, a pivotal intermediate in the biosynthesis of many carbon-containing compounds. Recently, Dr. David LaPorte (GM 33927) has completed some significant and fascinating work on the regulation of the glyoxalate bypass, and this report will highlight those findings.

The glyoxalate bypass enzymes, isocitrate lyase and malate synthase, are induced during growth of E. coli on acetate but are suppressed in the presence of many carbon sources such as glucose. The diversion of isocitrate through the glyoxalate pathway is regulated, in part, by the phosphorylation of isocitrate dehydrogenase (IDH), the Krebs cycle enzyme which competes with isocitrate lyase for a common substrate. Phosphorylation of IDH renders this enzyme inactive, decreasing the flux of isocitrate through the Krebs cycle and thus shunting some of it through the glyoxalate bypass instead. (In the cell some 20% of the IDH molecules do not get phosphorylated, so there is still some flow of intermediates through the Krebs cycle.)

Dr. LaPorte refers to the enzyme that inactivates IDH as IDH kinase/phosphatase, since he has accumulated a substantial amount of evidence that it is indeed bifunctional, catalyzing both the addition (kinase activity) and the removal (phosphatase activity) of the phosphate moiety. At the outset of the work Dr. LaPorte and co-workers found that the two activities co-purified and were associated with a protein (apparent molecular weight of 66,000) that appeared homogeneous by all the criteria that were applied. These data don't exclude the possibility that the two activities are associated with nonidentical subunits of a heterodimeric protein or with distinct proteins with very similar properties. To examine these possibilities a study of the gene(s) associated with these activities was undertaken. Physical and functional mapping of this cloned IDH kinase/phosphatase gene indicated that both the IDH phosphatase and IDH kinase activities are encoded by an 1,800-base pair sequence. This cloned sequence, when incorporated in a suitable expression vector, produced a protein with a molecular weight of 66,000---identical to that of the protein purified from acetate-fed E. coli. Since a protein of this size requires an 1,800-base pair coding sequence it was concluded that IDH kinase and IDH phosphatase are expressed from a single gene. This finding strongly suggests that both activities reside on the same polypeptide chain. The nucleotide sequence of the cloned gene was also established by Dr. LaPorte.

Other work carried out by Dr. LaPorte during this genetic study showed that the IDH kinase/phosphatase enzyme is expressed from the same operon that also codes for isocitrate lyase (aceA) and malate synthase (aceB). The gene for the IDH kinase/phosphatase activities, designated as aceK, was shown to be downstream from aceA; this latter gene was already known from the work of others to be downstream from aceB. It was, of course, quite satisfying to find these three genes, which have a common biological purpose, on the same operon.

The common motif among multifunctional proteins is one of structurally autonomous domains linked by a stretch of polypeptide, and it was originally suspected that IDH kinase/phosphatase would conform to this usual pattern. However, recent work in Dr. LaPorte's lab strongly indicates that the protein has a single active site that catalyzes both reactions in vivo. In deletion experiments, for example, Dr. LaPorte found that when 51 base pairs were deleted from the 3' end of aceK both activities were completely abolished. If the IDH kinase and phosphatase reactions are catalyzed by structurally separate domains, then this deletion would be expected to affect only one activity --- that associated with the C-terminus of the protein. It is possible, of course, that the deletion of the small C-terminal segment of

protein could severely disrupt the ability of the nascent protein to fold properly into a two domain structure; however, in most multifunctional/multidomain proteins this kind of behavior has not been observed.

The structural interdependence of the IDH kinase and phosphatase activities was also indicated by the properties of gene products of two recently isolated mutant alleles, aceK3 and aceK4. Both of these protein --- each resulting from a single point mutation --- showed significant alterations in their two activities. If the wild type enzyme had separate active sites for each of the activities, then a single point mutation would likely affect only one of the activities.

One other approach that Dr. LaPorte has taken to understanding the organization of the active site(s) of IDH kinase/phosphatase exploits the enzyme's nucleotide requirements. Dr. LaPorte has found that the IDH phosphatase activity is absolutely dependent on ATP or ADP. The kinase activity is, of course, also dependent on ATP, which acts as the phosphate donor. Examination of the amino acid sequence which was deduced from aceK for the IDH kinase/phosphatase protein revealed a single consensus ATP binding site. This sequence included the "invariant" lysine, which has been found in every protein kinase which has been sequenced to date. Using site-directed mutagenesis of aceK, this lysine was converted to a methionine, to produce a mutant protein in which both IDH kinase and IDH phosphatase activities were reduced to levels undetectable in vitro. Western blot analysis indicated that this loss of activity had not resulted from proteolysis of the mutant protein. These observations suggest that IDH kinase and IDH phosphatase share a common ATP binding site and that these reactions may well occur at the same active site.

IDH kinase/phosphatase, Dr. LaPorte has found, is also a very active ATPase. This was an important finding, since it rounded out a working model for this enzyme. According to this hypothesis, the kinase and phosphatase reactions occur at the same active site and the phosphatase reaction results from the kinase back-reaction tightly coupled to ATP hydrolysis. According to this model, the phosphatase reaction requires the formation of a ternary complex consisting of IDH kinase/phosphatase, phospho-IDH and ADP. (ADP could be bound directly or generated from ATP, by the ATPase activity.) The phosphate is then transferred from phospho-IDH to ADP (the back reaction of IDH kinase) and then to water (the ATPase reaction).

Bifunctional enzymes which are utilized in vivo for the catalysis of opposing metabolic reactions are quite unusual, but not unknown. An enzyme of this sort would be expected to be sensitive to certain metabolites that would shift the activity between the phosphatase and kinase mode. Indeed, Dr. LaPorte has found that a number of substances, including pyruvate, AMP, and 3-phosphoglycerate, activate IDH phosphatase and inhibit IDH kinase.

Dr. LaPorte's work on IDH kinase/pnosphatase is continuing and can be expected to add further definition to its interesting properties as a catalyst and as a metabolic regulator in E. coli.

Warren C. Jones

Warren C. Jones, Ph.D.

Memorandum

Date November 6, 1989

From Chief, Molecular Basis of Disease Section, CMBD, NIGMS

Subject Research Highlight: Aspartate Transcarbamylase (Dr. Howard Schachman, U. Cal., Berkeley, GM 12159; Dr. William Lipscomb, Harvard U., GM 06920; and Dr. Evan Kantrowitz, Boston College, GM 26237)

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Aspartate transcarbamylase (ATCase) catalyzes the first step in the biosynthesis of the pyrimidine nucleotides, compounds that are among the principal precursors for DNA biosynthesis in all organisms. The catalytic function of ATCase is to facilitate the reaction of carbamyl phosphate with aspartate to yield inorganic phosphate and N-carbamyl-aspartate (NCA); this latter material supplies all the atoms that comprise the pyrimidine skeleton of CTP and UTP. In addition to its catalytic function, ATCase controls the rate of the entire pyrimidine biosynthesis pathway, since formation of NCA is the slowest step in the pathway. However, the rate of formation of NCA can be raised or lowered by a combination of regulatory mechanisms. Since DNA biosynthesis is such a critical component of overall cellular regulation, it has been of considerable interest to understand how ATCase controls the rate of pyrimidine biosynthesis.

ATCase from E. coli is indeed a complex enzyme, being composed of twelve polypeptide chains of two types (see Figure 1). Each of the six larger, or catalytic, chains (C1 through C6) has a molecular weight of 33kD, and they are

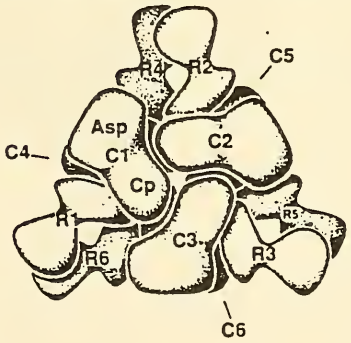


Figure 1. Schematic representation of the quatenary structure of ATCase. The 6 catalytic (C) and regulatory (R) subunits are numbered. Catalytic chains C1-C2-C3 and C4-C5-C6 constitute the catalytic trimers, while the regulatory chains R1-R6, R2-R4, and R3-R4 constitute the regulatory dimers. The catalytic chain is composed of the aspartate (Asp) and the carbamyl phosphate (Cp) domains.

grouped together in two trimers (Figure 1). Each of the smaller, or regulatory, chains (R1 through R6) has a molecular weight of 17kD, and they are organized as three dimers in the intact ATCase (Figure 1). The trimer, containing the catalytic sites, is active even in the absence of the

regulatory dimers. The dimeric regulatory subunits bind both purine and pyrimidine nucleotides, which serve in countervailing fashions to modulate the activity of ATCase. Thus, since ATP (a purine nucleotide) enhances the activity of the enzyme while CTP and UPT (pyrimidine nucleotides) inhibit the activity, the regulation of ATCase tends to maintain an intracellular balance of purine and pyrimidine nucleotides. In addition to this "heterotropic" control exerted by ATP, CTP and UTP, ATCase is also under "homotropic" control by its own substrates, carbamyl phosphate and aspartate. This latter phenomenon is manifest as homotropic cooperativity wherein the binding of aspartate, for example, to one active site in either of the catalytic trimers triggers a sharply enhanced affinity of all remaining subunits for substrates. This change in functional states of the enzyme (low affinity to high affinity) has been rationalized in terms of two different structural forms of the ACTase protein. Thus, it has been proposed on the basis of considerable data, including X-ray crystallography, that ACTase can exist in either a less active-low affinity structure (the T-state) or a more active-high affinity structure (the R-state) and that aspartate binding at one site in the T-state is effectively a switch that causes all remaining sites to adopt the R-state. These structural and functional properties of ATCase have been delineated over many years by Dr. Howard Schachman (GM 12159) and by Dr. William Lipscomb (GM 06920), working independently. This highlight will focus of the X-ray crystallographic work carried out on ATCase by Dr. Lipscomb and on more recent site-directed mutagenesis studies of this enzyme carried out by Dr. Evan Kantrowitz (GM 26237). These efforts have provided a working model at the molecular level that accounts for the enzyme's homotropic cooperativity.

The structure that Dr. Lipscomb has deduced for the T- and R-state of the catalytic subunits of ATCase are represented below--obviously in a highly schematic form. As is apparent from the figure, each catalytic subunit is

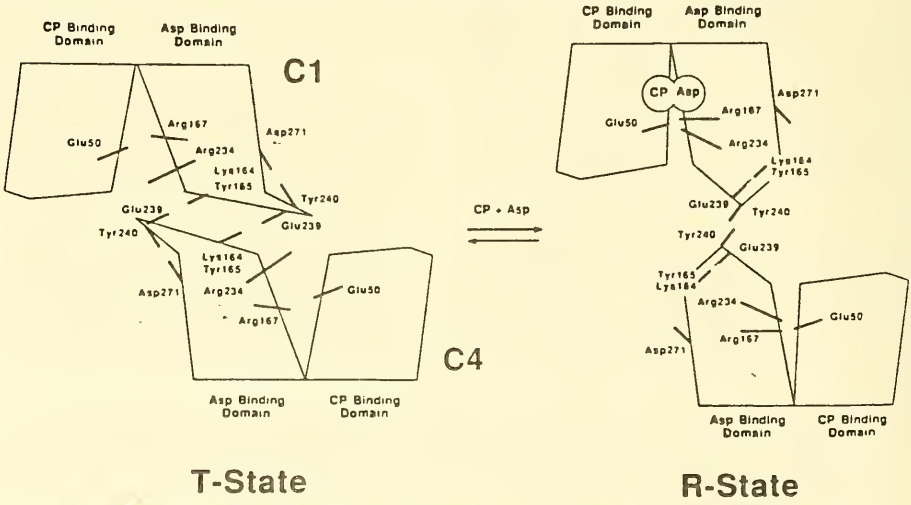


Figure 2. Model for the mechanism of homotropic cooperativity in ATCase.

composed of two domains: one which binds carbamyl phosphate and one which binds aspartate. In the T-state these domains are further apart than in the R-state; but, it appears from Dr. Lipscomb's work that as the substrates bind to a single subunit in the T-state, a structural reorganization occurs that causes the two domains of that subunit, C1, for example, to snuggle closer to one another (see figure). This reorganization of the target subunit causes a similar reorganization of a companion subunit (C4), thereby heightening its affinity for the substrates. The X-ray data indicate that there are numerous specific structural differences in the T and R states, and a few of the prominent differences have received significant attention. For example, the T-state shows an intrasubunit interaction between the polar side-chain groups of Tyr-240 and Asp-271 that could stabilize that structure; this interaction is absent in the R-state. On the other hand, the R-state shows an intrasubunit (but interdomain) interaction between the polar groups of Glu-50 and Arg residues at 167 and 234; this stabilizing interaction is absent in the T-state.

This understanding of homotropic cooperativity in ATCase has recently been given additional credence by the work of Dr. Kantrowitz. He and his coworkers acquired more detailed information about the function of the interactions involved in the two structural states of ATCase by using site-directed mutagenesis to perturb selected interactions. For example, when Dr. Kantrowitz replaced Tyr-240 with Phe, the mutant enzyme evinced a substantial reduction in cooperativity and a marked increase in affinity for aspartate; the mutant seemed to be predominently in an R-like state. Thus, it appears from this functional analysis of the mutant that the "lost" Tyr-240/Asp-271 interaction contributes to the stabilization of the T-form of the wild-type enzyme. In other work, Dr. Kantrowitz replaced Glu-50 with a Gln residue and found that the resulting mutant had low affinity for substrate, low catalytic activity, and no cooperativity. It thus appears that the Glu-50 interaction with Arg 167 and Arg 234 in the wild-type enzyme is critical not only to the formation of the high activity, high affinity catalytic site, but also to the maintenance of cooperativity.

These and other studies by Drs. Kantrowitz and Lipscomb have contributed substantially to the current understanding of homotropic cooperativity in ATCase. These studies also illustrate the value that the complementarity of two modern techniques have in enhancing the credibility and coherence of a research finding.

Warren C. Jones, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

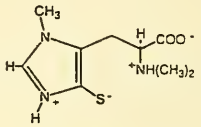
Memorandum

Date March 24, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "Ovothiols - a New Class of Anti-oxidants";
5 R01 GM23910-12 (Shapiro, Bennett), University of Washington, Seattle, WA

To Director, NIGMS
Through: Program Director, CMBD, NIGMS



Ovothiol C

Dr. Bennett Shapiro has discovered a new family of anti-oxidant compounds, the 1-methyl-4-mercapto-histidines. These aromatic thiol components, termed ovothiols (see diagram), are found in marine organisms and have unique chemical and biochemical properties which make them far more effective reducing agents than the aliphatic thiols such as glutathione.

Peroxides and oxygen radicals are potentially toxic in biologic systems but also can serve important functions. In mammalian cells an "oxidative burst" occurs in phagocytic cells, during which oxygen consumption is greatly increased, and hydrogen peroxide (H_2O_2) and superoxides, which aid in the destruction of microorganisms, are generated. In the sea urchin egg, a respiratory burst oxidase (long sought after and also recently discovered by Dr. Shapiro) gives rise to H_2O_2 which functions as an extracellular oxidant to cross-link the fertilization envelope. This structure is an extracellular protein coat which provides a physical block to polyspermy, and surrounds and protects the developing embryo until hatching (see Fig. 1).

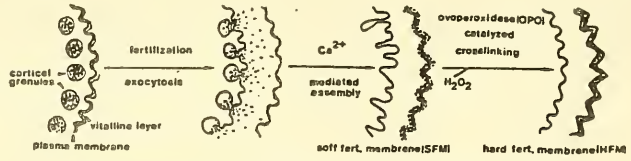


FIGURE 1. Cortical granule exocytosis and fertilization membrane assembly.

It is assembled from exocytosed egg cortical granule proteins and the extracellular vitelline layer, in a calcium-dependent process. One of the structural proteins is called proteoliasin, which mediates the incorporation of the cross-linking enzyme, ovoperoxidase, into the fertilization envelope. Ovoperoxidase then catalyzes the "hardening" of the envelope by the cross-linking of di-tyrosyl bonds in the presence of H_2O_2 (see Fig. 2). (Similar hardening of mouse egg zona pellucida, catalyzed by mammalian ovoperoxidase, also has been reported.)

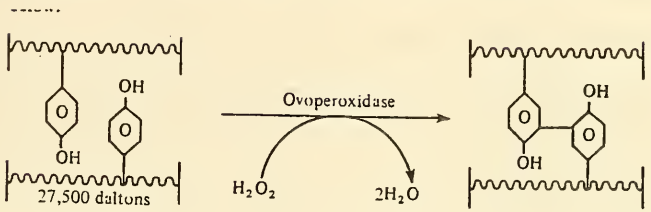


Figure 2

Once this important fertilization event is accomplished, the egg needs a way to inactivate the toxic H_2O_2. This is accomplished by the ovothiols, present in the reduced state in sea urchin eggs at a concentration of 5 mM. This non-enzymatic, aromatic ovothiol is 50 times more active than the aliphatic glutathione. Ovothiol C, with a pKa of 2.3 (instead of the 8.7 of glutathione) provides a striking intracellular environment for the sea urchin egg, which thus is 4-5 mM in thiolate ion. These compounds have been found in all invertebrate eggs and those of several fish species. They replace the enzyme glutathione peroxidase in eggs by serving the same function, i.e. detoxification of H_2O_2. The unique reactivity of the thiolate anion of ovothiol is associated with its very low pKa and its ability to react with oxygen centered radicals via one electron transfer reactions, where it is up to 10^4-fold favored over glutathione. The redox potential of oviothiol C is 84 mV positive with respect to glutathione; therefore in eggs it will always be in the reduced state.

Thus the newly discovered ovothiols represent a class of non-enzymatic, protective anti-oxidants which consume H_2O_2 and are then reduced (regenerated) by glutathione. This ovothiol replaces the function of glutathione peroxidase in eggs and also is more effective than the alternative egg catalase in destroying H_2O_2 at the high concentrations produced during fertilization. Dr. Shapiro is now beginning to study the role of ovothiols in mammalian cells where it may serve as an unusual anti-oxidant to provide protection from radical related damage, such as that which can occur during inflamation.

Marion M. Zatz, Ph.D.

1. Turner, Eric, Hager, Lisa J., and Shapiro, Bennett M., Science, vol. 242:939-941, Nov. 11, 1988.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 21, 1988

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT; "Regulation of cAMP-Dependent Protein Kinase"; GM32875-06 (McKnight, George S.), University of Washington, Seattle, Washington

To Director, NIGMS
Through: Program Director, CMBD, NIGMS

Cyclic AMP (cAMP) is an important cellular compound known to regulate metabolism in bacteria, and to act as a "second messenger" in eukaryotic cells where it plays a key role in regulating fundamental cellular processes such as division, growth, differentiation, secretion and transcriptional activity. Although the role of cAMP as an intracellular effector is conserved throughout the evolutionary scale, its mode of action in prokaryotes and higher organisms has diverged considerably. In bacteria, cAMP mainly binds to proteins which act as direct transcriptional modulators of cAMP-responsive genes. Eukaryotic cells contain a family of cAMP-dependent protein kinases which mediate most of the known cAMP "second messenger" effects. Thus the regulation of cAMP-dependent protein kinase synthesis and enzymatic activity is an important control point in the regulation of cellular responses.

The cAMP-dependent protein kinases are themselves complex, existing as a tetrameric holoenzyme which contains two catalytic subunits (C), and two regulatory subunits (R). Isoforms of C and R also exist, designated Cα and Cβ, and RI, RII, and RIII. Each of the R subunits is capable of binding two molecules of cAMP, which then results in dissociation of the R subunit dimer from the enzymatically inactive holoenzyme and release of the active C subunits. In order to probe the events which control enzyme assembly and function, Dr. McKnight has cloned two isoforms of the C subunit (Cα and Cβ) and has isolated mutant sequences of the RI subunit. Expression vectors encoding for Cα, Cβ or RI proteins were introduced into cells which then were studied for expression and activity of cAMP-dependent protein kinase.

In the first series of experiments[1], both 3T3 fibroblasts and AT-20 pituitary cells were transfected with Cα and Cβ plasmids. As expected, the transfectants expressed elevated cAMP-dependent protein kinase and associated enzymatic activity Surprisingly however, elevated levels of RI protein also were detected, although there was no change in RI mRNA. It is believed that this co-ordinate regulation of cAMP-dependent protein kinase subunits is due to a stabilization of the RI subunit protein by the C subunit in the holoenzyme complex; RI apparently is rapidly degraded in the absence of a C subunit. These results are consistent with the idea that cells normally produce an excess of RI subunit to prevent spontaneous activation of the holoenzyme. These studies also should permit a further examination of the respective roles of the Cα and Cβ isozymes, which are encoded by distinct genes, are tissue specific in expression, but share 91% amino acid sequence homology.

1. J. Biol. Chem., vol. 262, pp. 15202 - 15207, 1987
2. J. Biol. Chem., vol. 262, pp. 13111 - 13119, 1987

In the second series of experiments[2], expression vectors were constructed for mutant forms of RI, in which a specific amino acid was altered in the highly conserved cAMP binding domain of the RI protein. In one case a neutral glycine at position 200 was replaced by a glutamic acid residue, and in the second case a neutral glycine at position 324 was replaced by an aspartic acid residue. These neutral to acidic changes imply changes in electrostatic charge at the two cAMP-binding sites. In both cases the amino acid substitution resulted in an increased inhibition of cAMP-dependent kinase activity of 20-400 fold, depending on the amount of mutant RI protein expressed. In transfected adrenocortical tumor cells, this inhibition of enzyme activity was accompanied by resistance to growth inhibitory effects of cAMP and defective steroid synthesis.

In summary, Dr. McKnight is using a molecular genetic approach to introduce wild type and mutant catalytic and regulatory subunit genes into mammalian cells. This approach provides a unique tool to probe the assembly of the cAMP-dependent protein kinase holoenzyme, the regulation of enzymatic activity, and the functions of the individual isozymes in the cell. Further insights also can be expected as to the fine structural requirements for cAMP binding to R subunits. Ultimately this approach should yield a better understanding of the role of cAMP and protein phosphorylation in the process of intracellular signalling.

Marion Zatz

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 12, 1988

From Program Administrator, CMBD, NIGMS
Through: Program Director, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: Inhibitors of Serine Proteases - A Link to Alzheimer's Disease? Dennis Cunningham (GM31609-05), University of California, Irvine, California; Huntington Potter (GM35967-03), Harvard University, Boston, Massachusetts

To Director, NIGMS

Serine proteases occur widely in tissues and extracellular spaces as well as in the blood. Tissue-associated serine proteases are believed to play a major regulatory role in a wide array of cellular processes such as cell division and cell movements, while the free plasma proteases are integral components of coagulation, fibrinolysis and complement cascades. Dr. Cunningham's laboratory has identified, and studied extensively, a class of tissue serine protease inhibitors called protease nexins (PN's). These nexins appear to control the activity of serine proteases at or near cell surfaces, via interaction with the extracellular matrix, and to influence a number of important physiological and pathological events. The term nexin was originally used by Dr. Cunningham to signify the link between serine proteases and their inhibitors which resulted in extracellular matrix complexes. Three protease nexins have thus far been identified, termed PN-1, PN-2 and PN-3. PN-1 has been most extensively studied while PN-2 and PN-3 are in earlier stages of investigation.

PN-1 is secreted by a variety of cells, such as fibroblasts, muscle and epithelial cells and covalently binds in a stoichiometric 1:1 ratio with extracellular serine proteases at their serine catalytic site. The complexes bind back to the cell surface where they then are endocytosed and degraded intracellularly. At least seven serine proteases have been shown to be targets of PN-1, including thrombin, urokinase, plasmin, c'1r and c'1s complement components, plasma kallikrein and factor XII. These protease targets are the same as those for the c'1 protease inhibitor, which is found free in plasma and bears strong structural homologies to PN-1. Dr. Cunningham's most recent studies have demonstrated that PN-1 is bound both to the surface of the fibroblasts which secrete it, as well as to the extracellular matrix where it co-localizes with fibronectin; this binding is apparently mediated by a specific binding site on PN-1 for sulfated glycosaminoglycans (GAGS), such as heparan sulfate. These GAGS are present on the fibroblast cell surface and have been shown to accelerate the rate of complex formation between PN-1 and its target proteases. As an example of how

these complex interactions modulate cellular processes, Dr. Cunningham has shown previously that extracellular thrombin stimulates fibroblast division. He now has demonstrated that secretion of PN-1 by fibroblasts and binding of PN-1 to fibroblast surface heparan-sulfate molecules accelerates PN-1/thrombin complex formation, thereby regulating thrombin's mitotic influence. Another example of the fine tuning of physiological responses by nexins is provided by the glial cells in the brain, which secrete a related PN-1 molecule which promotes neurite extension by inhibiting the effect of thrombin on neurite retraction.

Recently Dr. Cunningham's lab has succeeded in purifying to homogeneity a second protease nexin, PN-2. Three targets for this nexin have thus far been identified; they are the α-subunit of nerve growth factor (NGF), epidermal growth factor (EGF) binding protein, and most recently chymotrypsin. The fact that PN-2 most strongly interacts with and inhibits chymotrypsin is of particular interest because of independent, but possibly related studies, which have come out of Dr. Potter's lab at Harvard. These studies have revealed that a serine protease inhibitor, α-1 anti-chymotrypsin, which occurs physiologically as an acute phase plasma serine protease, also is an integral component of the amyloid deposits in neuritic brain plaques characteristic for Alzheimer's disease. Neuritic plaques which appear in aged non-human primates also contain this serine protease inhibitor, thus providing a model system for further studies at the microscopic and molecular levels, of the role of serine proteases and their inhibitors in senile plaque formation.

This finding stems from the general problem under study in Dr. Potter's lab, namely regulation of expression of tissue specific-proteins. Dr. Potter's focus has been on the tissue specific-proteins of the immune and nervous systems, and in the case of Alzheimer's disease, those brain proteins that are aberrant in structure or amount. The amyloid proteins were an obvious choice to study since it has been known for some time that the brains of Alzheimer's disease patients have as a prominent pathological feature, neuritic plaques and neurofibrillary tangles containing amyloid protein deposits. While such histological features are found in brains of normal aging individuals, they are much more prevalent in the brains of Alzheimer's patients as well as in the brains of Down's syndrome individuals who have an extra copy of chromosome 21 and develop Alzheimer's disease at an early age. Since it is known that a significant component of Alzheimer's patients inherit the disease as an autosomal dominant genetic trait, several groups sought to look more closely at the proteins associated with the amyloid plaques and more specifically at the genes which might encode these proteins. Initially much excitement was generated by the finding that the gene for a major amyloid protein component, termed β-protein, was located on chromosome 21, since the association of Down's syndrome with Alzheimer's disease as well as family studies had implicated chromosome 21 as a likely site for the causative genetic lesion. While more recent work has shown that the gene for β-protein is not the site of the genetic lesion in families which show an inheritance pattern related to chromosome 21, other studies have suggested a role for protease inhibitors in modulating amyloid protein levels.

Dr. Potter and his group's approach led them to focus on one such protein in the amyloid plaques. Initially they produced antibodies to amyloid material

isolated from post-mortem brain tissue. The antisera were used to screen a cDNA human liver expression library. Three clones were identified which contained the cDNA for α-1 anti-chymotrypsin, a potent protease inhibitor known to be produced by liver and present in high concentrations in serum. It was then found that antibodies to purified α-1 anti-chymotrypsin stained the amyloid deposits in Alzheimer's brain tissue, as well as those in the brains of aging monkeys. A second group has also shown that two alternate forms of amyloid mRNA are produced in the brain. The longer form contains a 168 base pair insert which may be spliced out and encodes for a sequence that resembles a protease inhibitor. The finding of a serine-protease inhibitor associated with Alzheimer's lesions raises several interesting questions about the pathogenesis of the disease which now can be approached experimentally.

The first question is whether α-1 anti-chymotrypsin is transported to the brain lesions across an impaired blood-brain barrier from the serum, or produced locally. Initial studies indicate that mRNA for α-1 anti-chymotrypsin is present in low levels in the brain and thus could be produced in situ. A second question is what role (if any) α-1 anti-chymotrypsin plays in the development of amyloid deposits. Since amyloid is a complex association of multiple proteins with unique physical and chemical properties, α-1 anti-chymotrypsin could simply be another component of amyloid. Alternatively, this protease inhibitor could play a role in formation of amyloid deposits by several different mechanisms. Overexpression could be related to an inhibition of a protease which normally degrades the amyloid proteins, such as β-protein. Underproduction might allow excess activity of a protease which cleaves the β-protein from its precursor, permitting the release of the self-aggregating β-protein fragment. In this regard it is interesting to note that the cleavage sites of the precursor molecule known to generate β-protein fragments would be expected to be cleaved by a chymotrypsin-like protease. A third question is whether the location of the α-1 anti-chymotrypsin gene can shed any light on the genetic inheritance of Alzheimer's disease. Two loci to date, one on chromosome 21 and one possibly on chromosome 14, have been linked to inheritance of Alzheimer's disease in family studies. While the β-protein maps to chromosome 21, the α-1 anti-chymotrypsin gene is located on chromosome 14. Clearly additional genetic linkage studies will be needed to establish the gene(s) involved in the pathogenesis of Alzheimer's disease.

In summary, a variety of physiological processes are activated or regulated by serine proteases. The activation step generally involves proteolysis which is very specific to a few amino acids. These irreversible reactions are modulated by protease inhibitors which, by their ability to limit the extent, duration and site of protease action, add a second dimension of specificity to these regulatory systems. The protease nexins are inhibitors that act in the vicinity of the cell to control cellular processes, such as cell division and cell movement. They mediate interactions with the extracellular environment and are themselves regulated by cell surface molecules. The nexins are structurally and functionally related to the plasma protease inhibitors which control reactions such as coagulation, fibrinolysis and complement activation. Tantalizing evidence now suggests that a serine protease inhibitor normally found in plasma, α-1 anti-chymotrypsin, is also present in amyloid plaques of Alzheimer's brain tissue and may be produced locally.

These basic studies on regulatory proteases and their inhibitors may ultimately provide clues to the control of a variety of physiological and pathological events, including those related to the pathogenesis of the degenerative lesions of Alzheimer's disease.

Marion Zatz

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 21, 1988

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT UPDATE: "Serine Protease Inhibitors II. A Second Link to Alzheimer's Disease?" Dennis Cunningham (GM31609-06), University of California, Irvine, California

To Director, NIGMS
Through: Program Director, CMBD, NIGMS

In an earlier research highlight (1/12/88) Dr. D. Cunningham's work on two serine protease inhibitors, PN-1 and PN-2, was described. PN-1 was shown to play a role in regulating cellular activities, such as cell movement, which are mediated by serine proteases at the interface of the cell surface and its extracellular environment. In addition, PN-2 had recently been purified and one of its target proteases, chymotrypsin, had been identified. This latter finding was of particular interest since work in Dr. H. Potter's laboratory (GM35967, Harvard University) had just demonstrated that the serine protease inhibitor, α-1 anti-chymotrypsin, was an integral component of the amyloid deposits in the neuritic brain plaques characteristic of Alzheimer's disease (AD)[1]. Dr. Cunningham now has demonstrated that PN-1 also may play a role in the symptoms and lesions of AD.

[2]Earlier studies had shown that amongst its many activities, PN-1 stimulated neurite outgrowth, and that in fact PN-1 has an amino acid sequence identical to that of neurite promoting factor, which is produced by glial cells. The neurotrophic activity of PN-1 depends upon its inhibition of thrombin, a plasma protease which brings about neurite retraction. Alzheimer's disease is characterized by neuronal loss and the accumulation of plaques and tangles in selected brain regions. The plaques of this disease are characterized by a central amyloid core surrounded by degenerating neurites and reactive glia. Thus Dr. Cunningham proposed that reduced PN-1 activity might be associated with thrombin-mediated retraction and deterioration of neuronal processes. Recently he has indeed found that PN-1 activity was markedly reduced (13% of control) in the hippocampus, visual cortex and cingulate cortex regions of brains from AD patients. Two other serine protease inhibitors, anti-thrombin III, and heparin co-factor II, were not different from control in the same samples.

These results suggest possible models to explain some of the features of AD. Reduced biosynthesis of PN-1 in brains of AD patients could result in increased thrombin levels that in turn result in disrupted interactions among neurites and altered neurite morphology. (Thrombin also is known to inactivate another neurotrophic factor, acidic fibroblast growth factor.) Alternatively, reduced levels of PN-1 may be secondary to increased thrombin levels which complex PN-1 and reduce its steady state concentration. This latter possibility could result from a disruption of the blood-brain barrier,

suspected to occur in AD, subsequent escape of prothrombin from blood vessels, and conversion in the brain to thrombin. This latter mechanism is consistent with the observation that degenerating neurites are observed around the periphery of angiopathic blood vessels in AD. Regardless of the mechanism, at least two different serine protease inhibitors now have been found to be altered in brains of AD. It is hoped that these observations eventually will lead to a better understanding of the pathogenesis of this degenerative disorder of the elderly.

Marion M. Zatz, Ph.D.

1. Abraham, C.R., Selkoe, D.J. and Potter, H.; Cell 52, 487-501, 1988

2. Gurwitz, D. and Cunningham, D.; Proc. Nat'l. Acad. Sci. 1988, In press

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 19, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "From the Clinical to the Basic to the Clinical;" R01 GM23200-14 (Carson, Dennis), Scripps Clinic and Research Foundation, La Jolla, CA

To Director, NIGMS
Through: Program Director, CMBD, NIGMS Carr

Experiments of nature sometimes lead to basic discoveries, just as basic research ultimately leads to clinical interventions. In the early 1960's observations on inherited immunodeficiency led to the realization that the immune system was populated by two distinct kinds of lymphocytes - T-cells and B-cells. This insight led to an explosion in the understanding of how the immune system works. Dr. Carson's fundamental research into adenine metabolism in normal lymphocytes stemmed from his interest in the basis of adenosine deaminase (ADA) inherited immunodeficiency; these studies in turn have given rise to a novel drug therapy which holds promise for the treatment of chronic lymphocytic leukemia (CLL).

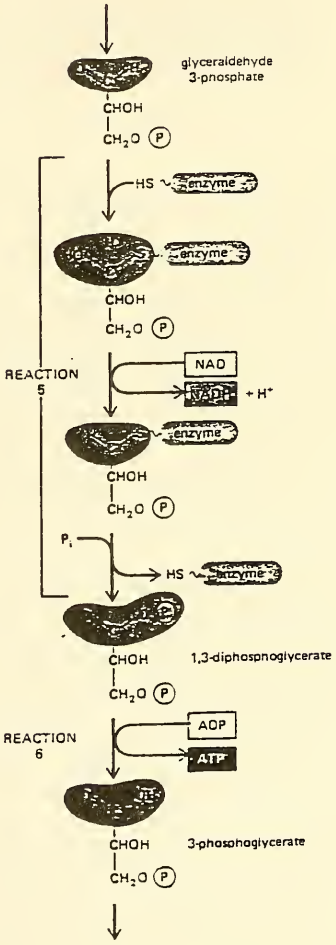
In ADA deficiency (see Highlight, 07/29/86) lymphocytes lack the adenosine deaminase enzyme which normally catabolizes deoxyadenosine (dAdo), a by-product of DNA precursor synthesis. This dAdo accumulates and becomes phosphorylated to dATP, which is an inhibitor of DNA synthesis. These toxic deoxyribonucleotides block the DNA repair of DNA strand breaks which normally occur even in non-dividing lymphocytes. The accumulation of DNA breaks in turn correlates with activation of a second enzyme, poly (ADP-ribose) polymerase (ADPRP) which uses NAD in poly-ADP ribose synthesis. Thus the cell's supply of NAD is severely depleted, which in turn limits the generation of ATP, an essential molecule of energy-dependent cellular processes (see Table 1 and Figure 1). Since lymphocytes have a limited capacity to synthesize NAD, and as much as 75% of NAD turnover is due to ADPRP consumption, the net result is a critical loss of cell functions and ultimate death.

TABLE 1. METABOLIC CHANGES IN dAdo-TREATED RESTING HUMAN LYMPHOCYTES

2 hr	↑	dATP
4 hr	↑	DNA strand breaks
	↑	poly(ADP-ribose) synthesis
	↓	RNA synthesis
8 hr	↓	NAD^+
24 hr	↓	ATP
48 hr		Loss of viability

Originally Dr. Carson set out to study ADA metabolism in an attempt to understand the molecular basis of ADA immunodeficiency. In the course of these studies however, he realized that the metabolic pathway triggered by DNA breaks might provide a mechanism for

Figure 1 Steps 5 and 6 of glycolysis: the oxidation of an aldehyde to a carboxylic acid is coupled to the formation of ATP and NADH (see also Figure 2–20). As shown, step 5 begins when the enzyme *glyceraldehyde 3-phosphate dehydrogenase* forms a covalent bond to the carbon carrying the aldehyde group on glyceraldehyde 3-phosphate. Next, hydrogen (as a hydride ion—a proton plus two electrons) is removed from the enzyme-linked aldehyde group in glyceraldehyde 3-phosphate and transferred to the important hydrogen carrier NAD^+ (see Figure 2–22). This oxidation step creates a sugar carbonyl group attached to the enzyme in a high-energy linkage. This linkage is then broken by a phosphate ion from solution, creating a high-energy sugar-phosphate bond instead. In these last two reactions, the enzyme has *coupled* the energetically favorable process of oxidizing an aldehyde to the energetically unfavorable formation of a high-energy phosphate bond. Finally, in step 6 of glycolysis, the newly created reactive phosphate group is transferred to ADP to form ATP, leaving a free carboxylic acid group on the oxidized sugar.



programmed cell death of damaged lymphocytes and could serve a physiological function in normal individuals. Taking this reasoning a step further, it might be possible to eliminate undesirable lymphocytes, such as the slowly dividing cells which accumulate in CLL, by treating patients with dAdo or an analog. The toxic effects of several ADA-resistant analogs were tested on lymphoid cell lines. (Since normal cells contain ADA, it was essential that the analogs be ADA-resistant.) One analog, 2-chlorodeoxyadenosine (2-ClAdo) proved to be very effective in selectively killing lymphocytes. Its mechanism of action is unique compared to drugs which are used in the treatment of acute leukemia, which act on the basis of inhibition of division of the rapidly proliferating leukemia cells. In CLL, the diseased cells are not rapidly dividing but rather seem to accumulate due to a long life span. In a small clinical trial (N=18) 55% of the CLL patients showed reduction of CLL cells and clinical improvement, without toxicity to bone marrow or other organ systems. The absence of

toxicity is related to the selectivity of 2-CldAdo for lymphocytes. An interesting additional benefit to the patients was the improvement in the autoimmune hemolytic anemia which often accompanies CLL. This second benefit is probably attributable to the destruction of autoimmune lymphocytes.

Thus the study of ADA immunodeficiency, one of the experiments of nature, has led not only to an understanding of the disease lesion, but to normal lymphocyte purine metabolism, and a potential rational therapy for CLL.

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 17, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlight: "Structure and Function of Plant Cell Wall Extensins"
R01 GM18639 (Staehelin, L.A.), University of Colorado

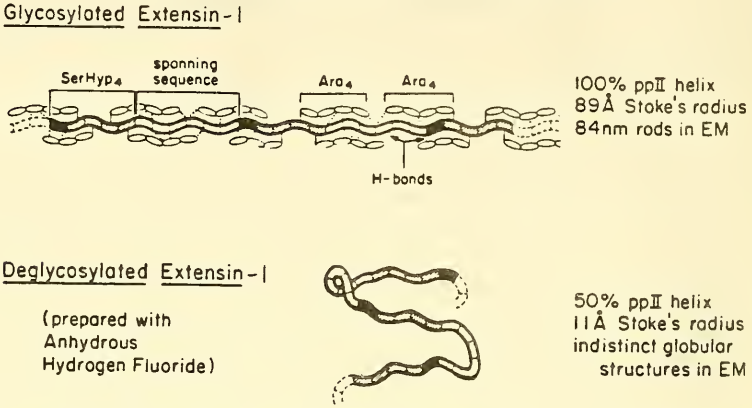
To Director, NIGMS
Through: Director, CMBD Program, NIGMS CAM

Even though the patterns of plant cell growth and development depend on the synthesis, secretion, and assembly of cell wall molecules, plant cell walls are often viewed as simple "wooden boxes" that provide mechanical protection and a buffered environment to the protoplast. This notion grew from the observation that plant cell walls are composed largely of polysaccharides, cellulose fibrils, and matrix materials; there is very little protein present. This simplistic concept has been challenged by Dr. Peter Albersheim's recent work at the University of Colorado, Boulder, demonstrating the chemical complexity of plant cell wall carbohydrates and their relationship to many overlooked physiological functions of the walls. He has provided descriptions of small (7 to 15 residue) oligosaccharide hormones, stored in cryptic form in the wall, which act in picomolar concentrations as signalling and regulatory molecules. These hormones control patterns of growth; orchestrate responses to environmental stresses such as invading microorganisms, chewing insects, and tissue wounding; and regulate developmental processes such as root formation, shoot formation, and flowering.

Dr. Albersheim's work has tremendous implications for both animal and plant research. It stimulated NIGMS grantee Dr. L. Andrew Staehelin (University of Colorado, Boulder) to undertake a systematic study of the biosynthesis, ultrastructure, and biochemistry of plant cell wall protein species, in order to understand the cell biology of the extracellularly-stored oligosaccharide hormones. His model system is the carrot, which grows readily in suspension cell culture and which produces somatic embryoes that differentiate in a characterized and reproducible manner. The dominant protein constituent of the primary cell wall of carrots (and all dicotyledonous plants) has been reported to be extensin. Extensin is an hydroxyproline-rich (up to 45 percent) glycoprotein. Two-thirds of its mass is carbohydrate, largely hydroxyproline-linked tetra-arabinoside but with some serine-linked mono-galactoside. Dr. Staehelin developed a combination of acid-urea PAGE and Sephacryl column chromatographic procedures to purify extensin monomers to homogeneity and determine their amino acid sequence. He discovered that extensin is actually two different, but related, protein species: The major protein is extensin-1, an 86kD molecule. By a combination of circular dichroism spectroscopy, measurement of the Stoke's radius, and digitized platinum-shadowed electron microscopy, Dr. Staehelin was able to show that the native protein exists as an elongated rod in an extended polyproline II

helical conformation. The molecule contains "kinks" at locations which correspond to unusual isodityrosine (IDT) residues (formed by peroxidation of tyr-lys-tyr sequences by peroxidase). Dr. Staehelin demonstrated that the carbohydrate moieties are essential for maintaining the extended conformation of extensin-1. Deglycosylation resulted in a globular structure:

ROLE OF CARBOHYDRATE IN EXTENSIN STRUCTURE



Model depicting the role of hydroxyproline-linked arabinosides in maintaining extensin in an elongated conformation.

Extensin-2 is a smaller (43kD), less kinked, and less hydroxyproline-rich (28%) glycoprotein, with only 1/3 the tyrosine content of extensin-1.

Dr. Staehelin prepared antiserum against each extensin species and used immunoelectron microscopy to examine their biosynthesis and in vivo structure. Extensin-1 monomers were secreted in the wall, co-localized with cellulose-containing regions but absent from the middle lamella and three-cell junction. Wall-bound peroxidases then covalently cross-linked the monomers by IDT residues; the intermolecular cross-links were preferentially formed at the ends of molecules. The cross-linked matrix was rigid and highly insoluble, and appeared to be woven around the cellulose fibrils, conferring strength to the cell wall. Expression of the extensin matrix was correlated with an inhibition of cell elongation, and increased resistance to pathogens. These functions depended on full glycosylation of the extensin-1 molecule and were inhibited by drugs which prevented arabinosylation. It appears that the intermolecular terminal cross-linking sites must be widely separated to form a matrix which functions properly in vivo, and that molecules must be fully glycosylated (as postulated in Dr. Staehelin's structural model) to achieve this separation.

Extensin-2 did not form oligomers *in vivo* or *in vitro* and mixed extensin oligomers were never observed. Therefore, it seems that extensin-2 is not cross-linked in the cell wall. Extensin-2 was preferentially distributed in the middle lamellae and corners, so its immunolocalization is complementary to that of extensin-1. It is possible that extensin-2 may be involved in disease resistance, since it has carbohydrate binding properties similar to the solanaceous lectins.

Dr. Staehelin will next turn his attention to the regulation of extensin synthesis to determine how they are assembled in cell walls during development and in response to pathogens and other stresses. His work should continue to give us fresh insight into the extremely interesting and dynamic plant cell wall structure and function.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Program Administrator, CMBD, NIGMS
Through: Program Director, CMBD, NIGMS

Subject Research Bullet: In Vitro Assay for Studying Force Generating Properties of Flagellar Dynein. Dr. Witman GM 30626-06 and Dr. Vallee GM 26701-09

To Director, NIGMS

The inner and outer arm of the flagellar axoneme generate the forces needed for flagellar movement in a variety of organisms. These flagellar arms contain ATPases called dyneins which are believed to be the force generating molecules in flagellar as is acto-myosin in muscle. Until recently there had been no method for studying this mechanochemical transducing activity (force producing properties) of isolated dyneins. Dr. George Witman and Dr. Richard Vallee have in a collaborative effort, presented data, (Nature, 17, December, 1987) indicating that axoneme dyneins can be studied in an in vitro assay as has been the case with myosin (Sheetz and Spudich reported as a research highlight 1984, bullet 1985). Prior to the Witman-Vallee joint effort, Dr Vallees' laboratory reported that the brain microtubule associated protein (MAP) 1 C is a microtubule activated ATPase with the structural and force producing properties of dynein. (MAP 1) C translocates microtubules in an in vitro gliding assay. Building on these experiments, the Vallee and Witman groups have shown that outer arm dynein isolated from sea urchin (Strongylocentrotus purpuratus) sperm can translocate calf-brain microtubles along the surface of a coverslip. These results demonstrate that outer arm dynein by itself is capable of generating shearing forces. This in vitro assay should greatly facilitate studies of the mechanisim of action of this important mechanochemical transducing dynein from diverse organisms and for elucidating the functions of cellular movement in flagella of many species.

Artrice V. Bader, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 17, 1988

From Program Administrator, CMBD, NIGMS

Subject Research Highlight: Dr. Richard Vallee, GM26701 MAP-1C: A Microtubule Translocator

To

Director, NIGMS
Through: Director, CMBD Program, NIGMS

Cytoplasmic microtubules are composed of tubulin and a number of microtubule associated proteins (MAPs) which are thought to be involved in microtubule function and in regulation of microtubule assembly. There are two classes of these microtubule associated proteins, high molecular weight and tau proteins. The high molecular weight MAPs have molecular weights of 200 kilodaltons (kD) to 300 kD, the tau proteins 60 to 70 kD. These accessory proteins are known to enhance the polymerization of brain microtubules. Dr. Richard Vallee's latest findings have created a new approach and direction for workers in this area. He began working on microtubules as a post-doctoral fellow in the laboratory of Dr. Gary Borisy, and subsequently has focused his research on MAPs. In 1984 and 1985 he reported, with Dr. Bloom, that the most prominent MAPs in brain tissue are a class of five high molecular mass proteins which they termed MAP 1A, MAP 1B, MAP 1C, MAP 2A and MAP 2B. Although some structural similarities exist between MAP 1A and MAP 1B, the MAP 1 polypeptides are less closely related than the MAP 2 species as judged by both biochemical and immunological analysis. MAP 2A and 2B appear to be related to each other in primary structure, while MAP 1C has consistently shown distinct properties from the other MAP species.

When these high molecular weight MAP's were first identified, they excited researchers in the cell motility field as possibly being candidates for a role in microtubule associated motility in the cell. Suggestions for this role were based on their similar subunit size to axonemal dynein, an ATPase which has been shown to provide the motive force for microtubule-associated motility in flagellar and ciliary axonemes. However, despite extensive biochemical characterization of partially purified MAP preparations, it has only been recently that a similar role for MAPs in motility has been realized.

In the search for a mechanochemical function of the high molecular weight proteins, MAP 1C has emerged as the front runner. Previously it was the least prominent of the high molecular mass species isolated from either gray or white matter microtubule preparations. However, high yields of MAP 1C were obtained when microtubules were prepared in the absence of added nucleotide. Using nucleotides for extraction it was discovered that ATP but not GTP could be used to extract MAP 1C from the microtubules. The ATP nucleotide extraction procedures provided not only a purer fraction but indicated that MAP 1C is unique among the high molecular MAPs and does bind to microtubules in a nucleotide dependent fashion.

In view of the ATP dependence of microtubule binding and the microtubule - activated ATPase activity of MAP 1C, Dr. Vallee and group designed experiments to further test the possibility of a mechanochemical function for this protein.

The exciting results of these experiments show that the MAP 1C ATPase is capable of translocating microtubules in an in vitro assay. MAP 1C properties in these experiments appear to be distinct from all of the other known high molecular weight MAPs, as well as the force-generating proteins kinesin and dynein, although some similarities to the latter two are noted. MAP 1C has several similarities with flagellar and ciliary dynein in sedimentation properties, effects of ATP binding, and, most striking, its response to ultraviolet light in the presence of vanadate and ATP. Such exposure results in specific limited cleavage of the molecule. Prior to this finding for MAP 1C, this clevage had been reported as being specific for dynein. Despite the similarities, differences between MAP 1C and dynein include the isolation of MAP 1C from the cytosolic fraction of brain homogenates; a lower ATPase activity and stimulation of the MAP 1C ATPase occurs at approximately a 100-fold lower concentration of microtubules.

In an in vitro assay utilizing axonemes on coverslips and MAP 1C-preparations, microtubule motility in the form of gliding was observed.

(Figure 1)

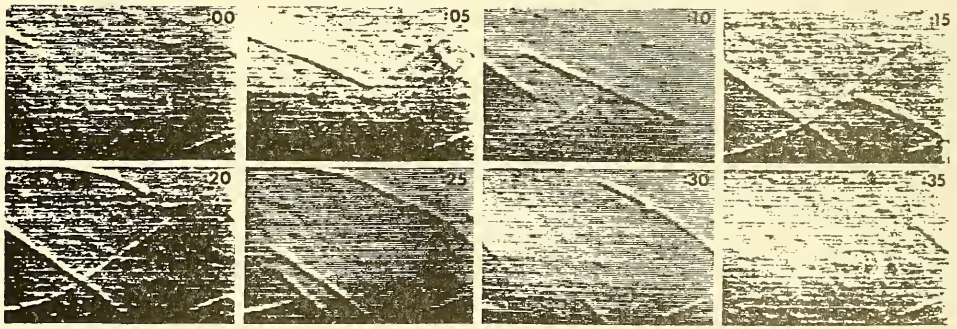


Figure 7. Motility of microtubules mediated by purified MAP 1C. MAP 1C was purified by sucrose density gradient centrifugation. The peak fractions were pooled (gel is shown in Fig. 8), dialysed into Tris/KCl buffer, and adsorbed to a coverslip. Microtubules were applied in the presence of 1 mM added MgATP. Most of the microtubules in the preparation showed gliding motility. The sequence presented here shows a typical field in the preparation. Seven microtubules can be observed to cross the field in the 35-s interval shown. One immobile microtubule may be seen at the lower right of the field. The sequence shown was taken 3 h after the preparation had been mounted for microscopic examination, and the average rate of movement had decreased from 1.25 (Fig. 9) to 0.97 μm/s. Bar, 2 μm.

Again the similarity to dynein is noted as axonemal dynein mediates microtubule gliding. By analogy with axonemal dynein, MAP 1C could be involved in the reorganization of cytoplasmic microtubules.

After MAP 1C and kinesin separately were added to coverslips and allowed to adsorb, axonemes were assayed for gliding motility in the presence of ATP. No movement was observed. When axonemes were applied to the coverslips containing previously adsorbed MAP 1C, axonemal gliding was observed and was always from the proximal (-) to the distal (+) end. In contrast, when added to coverslips with kinesin, gliding occured in the opposite direction.

(Figures 2 and 3)

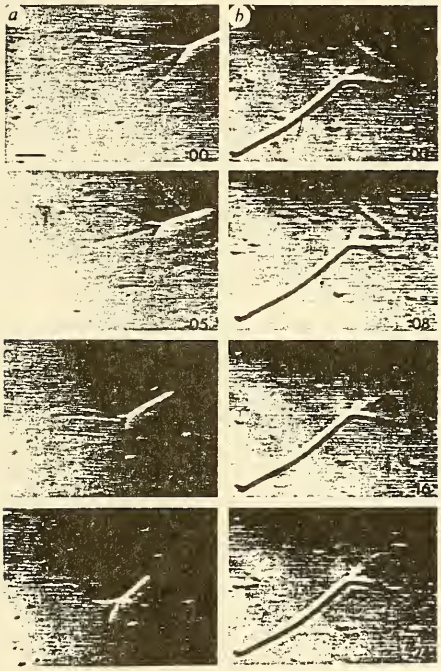


Fig. 2 Polarity determined using microtubules assembled onto flagellar axonemes. *a*, MAP 1C; *b*, kinesin. Purified calf brain tubulin[8] was allowed to assemble onto *Chlamydomonas* axonemes. As reported previously[10-13], the microtubules at the frayed (distal or +) end of the axoneme were longer than those at the compact (proximal or −) end. *a*, Gliding motility of the entire microtubule/axoneme complex. Short microtubules seen at the proximal end. *b*, Axoneme remaining stationary: microtubules seen only at the distal end and showed writhing movement, characteristic of microtubules experiencing a force but tethered at one end[6,14].
Methods. The motility assay and preparative methods were as described in Fig. 1. Purified tubulin[8] was added to a final concentration of about 1.5 mg ml^{-1} to flagellar axonemes in Tris/KCl buffer + 1 mM GTP and incubated for 5 min at 37 °C. Tubulin was added to the same final concentration to the sample of MAP 1C or kinesin already applied to the coverslip, and the axoneme/microtubule complex was then added and monitored for motility.

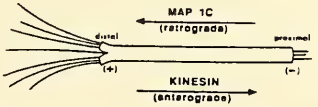


Fig. 3 Diagrammatic representation of force production. A *Chlamydomonas* axoneme is shown with long microtubules assembled onto frayed (distal or +) end, and short microtubules assembled onto compact (proximal or −) end[10-13]. Axonemes were driven towards their distal ends by MAP 1C (Figs 1, 2). Thus, a cytoplasmic structure attached to MAP 1C would be driven in the opposite direction, or towards the proximal end of the microtubule. In axons this would correspond to retrograde motility, as axonal microtubules have the same polarity as those in the *Chlamydomonas* axoneme relative to the cell body[24,25]. Axonemes and microtubules assembled at their ends were driven toward the proximal end of the axoneme by kinesin. This would correspond to anterograde motility for an axonal organelle.

These findings are consistent with the earlier data showing similarities of MAP 1C to dynein. Dynein is thought to produce force in a direction opposite to that observed for kinesin. There are those who believe this is consistent with the hypothesis that MAP 1C is a soluble form of dynein, and is responsible for retrograde organelle transport in cells. Drs. Paschal and Vallee conclude that in these studies with Chlamydomonas axonemes, MAP 1C acts as a retrograde translocator and causes movement in a direction opposite to the movement caused by kinesin. They feel that in this system kinesin is functioning as an anterograde translocator.

In conclusion, MAP 1C, the fifth of the high molecular mass MAPs to be described in detail, appears to be totally distinct from other proteins of this class and most importantly, has properties of a mechanochemical enzyme. Its biological function is unknown but MAP 1C, like kinesin, may play a role in transport of vesicular organelles with a retrograde function being currently favored. Also because of its similarities to axonemal dynein of flagella and cilia, which mediates microtubule sliding, it could be involved in the reorganization of cytoplasmic microtubules. This latter type of behavior has not been extensively described by Dr. Vallee's group but they feel that MAP 1C could be involved in the growth or the reorganization of the neuronal cytoskeleton. These findings have had a fundamental impact on the microtubule field and understanding the function of MAP 1C should provide further insight into the workings of neuronal cells and perhaps, other types of cells as well. Dr. Vallee and co-workers continue their quest and remain at the forefront of microtubule associated protein research.

Artrice V. Bader

Artrice V. Bader, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Deputy Director, CMBD, NIGMS

Subject Research Bullet: GM 36277-02 and GM 30387-06 (James Spudich) Stanford

To Director, NIGMS
Through: Director, CMBD, NIGMS

The movement of myosin molecules along actin filaments provides the driving force for a wide variety of cellular processes, including muscle contraction, cytokinesis, cytoplasmic streaming, and cell locomotion. The sliding filament model has long provided the theoretical basis for the shortening of sarcomeres in striated muscle, and variations on this theme have been proposed to account for many other myosin-based forms of contractility. The study of such mechanisms has come mainly from experiments using striated muscle fibers, in which the thick (myosin) and thin (actin) filaments are in precise array. Moreover, extracts of such muscle fibers have been shown to retain the ATPase and motile properties, allowing studies of the biochemistry in considerable detail.

Recently, however, work by NIGMS grantees using model systems from cellular systems from primitive eucaryotes promises to lead the way in our understanding of the molecular mechanisms of force generation. A couple of years ago researchers succeeded in preparing myosin-coated polystyrene beads which could be made to move along the actin cables extracted from the alga Nitella. This provided a fine assay of movement kinetics. Further refinement has allowed the attachment of myosin fragments, such as a short piece of heavy meromyosin, to be anchored, through antibody linkages, to fixed bacterial cells. This allows movement of the bacteria along the actin filament to be monitored. Most recently, investigators have succeeded in fixing myosin fragments from the slime mold, Dictyostelium, to coated coverslips. The movement of flourescently-labelled actin over this surface can then be measured. These in vitro assays are providing a great deal of important information about actin-myosin interactions. Recent findings, for example, provide strong evidence that myosin heads are extremely flexible. Other data are shedding light on the question of the spacing of reactive sites along the actin molecule. The early results are exciting, and show that systems using myosin from slime molds and actin from algae have much to contribute to our understanding of muscle contraction

Bert I. Shapiro, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 28, 1988

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "MCPs Forward; About Face"
R01 GM 19559-16 (Parkinson, John S.), University of Utah

To Director, NIGMS
Through: Program Director, CMBD, NIGMS

All cells, from the simplest bacteria to those making up complex mammalian tissues, need to sense changes in their microenvironment and to respond appropriately. Important examples of sensory transduction include cellular chemotaxis, reaction to changes in osmolarity and light, and stimulus-response coupling for hormones and growth factors. Dr. Parkinson has been studying the chemotactic responses of E. coli as a model system to dissect the molecular requirements for sensory transduction. Using mutants which he has constructed, Dr. Parkinson has been able to map structural domains of membrane transducing molecules and to correlate changes in individual amino acid residues with functional changes in the sensory transducers. He proposes that these primary structural changes translate into functional changes by altering the conformational dynamics of the transducer molecules.

E. coli swims by rotating its flagellar filaments. A counter clockwise (CCW) rotation produces forward motion, whereas a clockwise (CW) rotation results in turns and tumbles. The directed movement of E. coli is a chemotactic response elicited by chemical changes in its microenvironment, and is mediated by a family of membrane proteins known as methyl-accepting chemotaxis proteins (MCPs). These MCPs have been studied extensively by several groups including those of D. Koshland, J. Adler and G. Hazelbauer. The MCPs function by modulating the rotational behavior of the flagellar motor. Four different MCPs have been identified, TAR, TSR, TRG, and TAP. Each one monitors a different set of chemical stimuli through specific ligand-binding (chemoreceptor) sites displayed on the periplasmic side of the membrane (TAR monitors changes in aspartate and maltose; TSR monitors changes in serine, alanine and glycine; TRG and TAP monitor changes in sugars via an intermediate binding protein). Changes in receptor occupancy appear to induce conformational changes which are transmitted across the membrane to the cytoplasmic portion of the transducer molecule, which then interacts with a family of soluble chemotaxis proteins (CheA, CheB, CheR, CheW, CheY, and CheZ) to generate signals that control flagellar rotation. A feedback loop, catalyzed by the CheB and CheR enzymes, results in addition or removal of methyl groups at specific glutamic acid residues on the MCPs, canceling the excitatory signals and bringing about sensory adaptation which enables the cells to respond to further changes in chemoeffector levels.

The MCP transducer proteins contain approximately 550 amino acids, organized into two discrete structural domains (see Fig. 1) The N-terminal half comprises the periplasmic receptor domain, which is flanked at either end by membrane-spanning segments (TM1 and TM2). The C-terminal half of the molecule

comprises the cytoplasmic signaling domain, which also is flanked at each end by segments (Kl and Rl) which contain the methylation sites. These two domains are joined by a "linker" segment. The cytoplasmic portion of the MCP molecule is highly conserved, in contrast to the chemoreceptor domain and "linker" segment, which vary considerably in primary structure.

In order to dissect the roles of these structural components in transducer function, Dr. Parkinson generated and characterized three classes of MCP mutants with specific defects in ligand binding, transmembrane signaling, or sensory adaptation. Based on initial data, Dr. Parkinson has proposed a model whereby transmembrane signaling by MCP molecules is generated via direct transmission of conformational changes in the periplasmic receptor to the cytoplasmic signaling domain, through the linker segment. This model contrasts with alternate signal transduction mechanisms in which receptor occupancy results in ligand-dependent changes in receptor subunit aggregation state. The experimental approach to the structure-function analysis of sensory transducers was as follows: Plasmids carrying the structural genes for TAR or TSR were subjected to chemical mutagenesis. Transducer defect mutants were isolated by screening transformants for altered chemotactic behavior. These transducer mutations were mapped to their coding regions and appropriate oligonucleotide primers were used to determine the DNA sequence changes in the mutants. The primary amino acid sequence was then deduced from the nucleotide sequence.

The TAR transducer was chosen for study in order to define receptor site structure and initial events of transmembrane signalling. TAR recognizes two distinct chemoeffectors, aspartate which binds directly to TAR, and maltose, which is detected via a maltose-binding protein intermediate. The two binding sites are separate and it therefore is possible to detect mutants defective in aspartate binding and the aspartate chemotactic response, but which maintain signal transduction capacity for maltose. TAR mutants were identified which detected aspartate at a much reduced level of sensitivity, secondary to reduced affinity of the ligand for the receptor. Such mutants responded normally to maltose, thus confirming the specificity of the alterations to the aspartate binding site. All mutational changes could be attributed to three closely spaced codons which specify arginine residues at positions 64, 69 and 73; in each case the charged arginines were replaced by an uncharged amino acid. It is of interest to note that the TSR transducer, which also directly binds amino acids, and whose N-terminal primary sequence in general varies considerably from TAR, nonetheless has arginine residues at the same positions as does TAR. In contrast, TRG and TAP, which do not directly sense amino acids, have little sequence homology to TAR in this segment of the periplasmic domain. Since it is known that conformational changes occur in MCPs upon ligand binding, it is likely that receptor occupancy triggers conformational shifts in this highly conserved region which then are propagated to the signaling domain to modulate activity. Dr. Parkinson speculates that the arginine segment represents a major control point in the receptor domain; the ligand-free receptor might be strained by the electrostatic repulsion of the charged arginine residues, while receptor occupancy would lead to a less strained conformation.

A similar experimental approach was used to explore the structural features of the cytoplasmic signaling domain using TSR mutants. Fifty eight

non-chemotactic mutants were isolated which were locked into either a CW or CCW signaling pattern, indicating that MCP molecules may alternate between two active signaling states. All mutations were either in the "linker" segments or in the middle of the cytoplasmic domain. The cytoplasmic domain mutants involved residues identical in all four MCP transducers. Amino acid replacements at residue 394 or 407 resulted in mutants locked in either mode, suggesting that this region may act as a conformational "hinge" to control transitions between the CW and CCW signaling states. Mutations in the N-terminal half on the cytoplasmic domain conferred a CCW signaling bias whereas mutations in the C-terminal half of the cytoplasmic domain (residues 392 - 466) conferred a CW signaling bias. These locked mutants by definition could not switch from one active signaling state to another and also were therefore refractory to sensory adaptation. It is believed that a conformational change controlled by methylation state results in switching from CW to CCW, giving rise to sensory adaption , and that these closely linked mutations lock the MCP into one or another conformational and rotational signaling mode. Consistent with this notion was the finding that CCW mutants had high methylation states and CW mutants had low methylation states. The methylation levels ranged from much higher to much lower than wild type levels; these anomalous patterns probably reflect uncompensated control of the sensory adaption machinery by the CheB and CheR enzymes.

Dr. Parkinson identified additional TSR mutants with less extreme CW or CCW biased signaling properties, which were obtained as phenotypic suppressors of CheB or CheR mutants. Unlike the "locked" transducer molecules, these MCPs still respond to the sensory adaptation system and retain TSR function. These mutations clustered mostly in the linker segment and near the methylation sites in the cytoplasmic domain. Dr. Parkinson's work and that of Koshland[1] and others, indicate that MCP transducers mediate transmembrane signaling by direct transmission of induced conformational changes. It seems likely that the TM1 and TM2 spanning segments may interact with each other to maintain the conformation of the receptor domain, whereas the activity of the signaling domain is regulated by the interactions of the methylation segments. His model (Fig. 2) for sensory signaling is that receptor occupancy may propagate a conformational change to the signaling domain through relative shifts in the alignment of TM1 and TM2, leading to a push or pull on the linker segment; this in turn influences the conformation of the signaling domain, possibly by altering the relative alignment of the methylation segments.

Clearly much still needs to be learned before all the stages from ligand binding to sensory signaling are understood. But Dr. Parkinson has taken the first few steps in understanding how the primary sequence at the single amino acid level can contribute to the three-dimensional structure and function of sensory transducers.

Marion Zatz

[1]Falke, J.J. and Koshland, D.E., Jr., Science 237:1596-1600, 1987

Figure 1

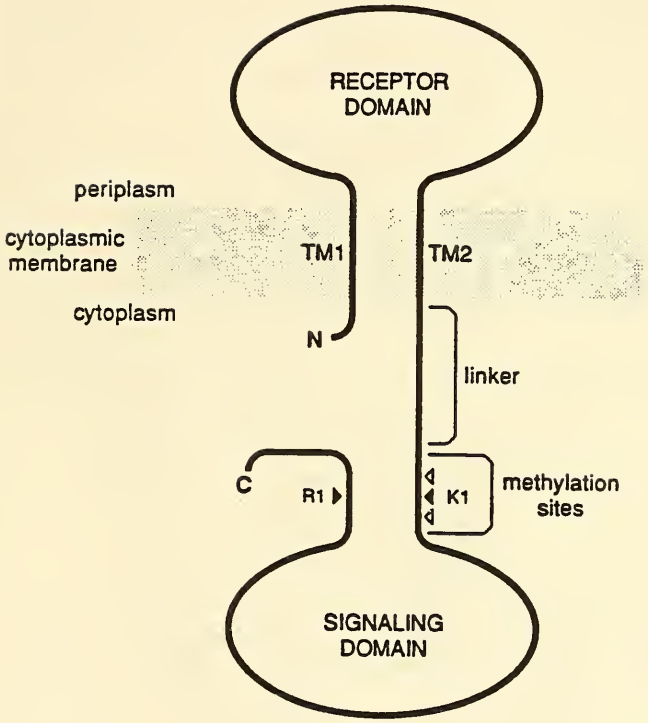
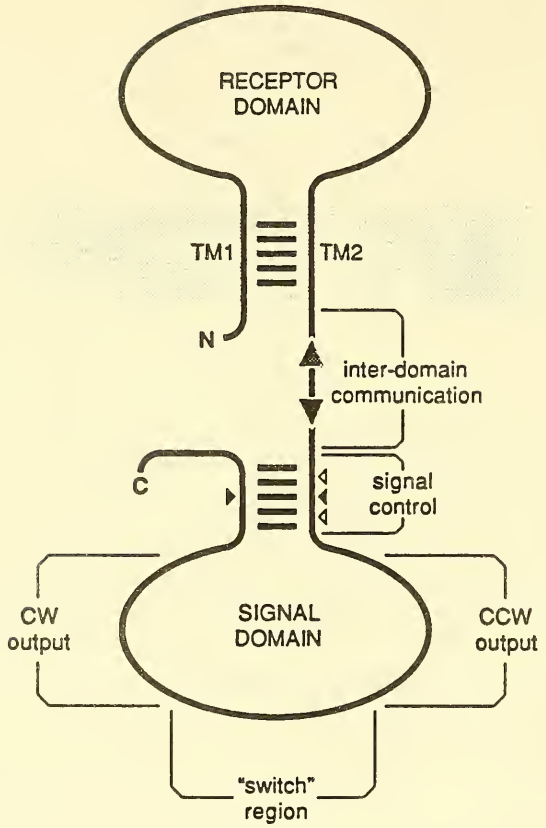

RECEPTOR
DOMAIN
periplasm
cytoplasmic
membrane
cytoplasm
TM1
TM2
N
linker
C
R1
K1
methylation
sites
SIGNALING
DOMAIN

Figure 2


RECEPTOR
DOMAIN
TM1
TM2
N
inter-domain
communication
C
signal
control
CW
output
SIGNAL
DOMAIN
CCW
output
"switch"
region

Memorandum

Date July 7, 1989

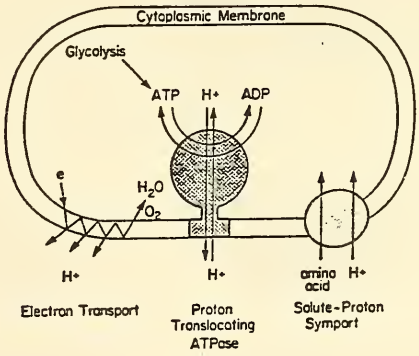
From Program Administrator, CMBD Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Assembly of the Escherichia coli Proton-Translocating ATPase" R01 GM 18539-19 (Simoni, R.D.), Stanford University

To Director, NIGMS
Through: Director, CMBD Program, NIGMS Excellent work

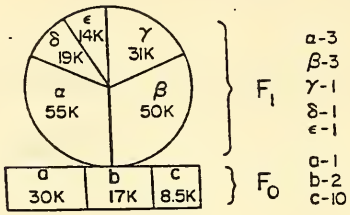
The F_1F_0 proton-translocating ATPase has been a focus of biochemical study for over 30 years. It is the central component in cellular energy transduction in all biological systems. (see Fig. 1) In mitochondria and in aerobic bacteria, this ATPase generates an electrochemical gradient of protons to carry out oxidative phosphorylation. In anaerobic bacteria, F_1F_0 hydrolyzes ATP to form a proton gradient essential for all other energy-dependent processes. And in chloroplasts, the ATPase is responsible for photophosphorylation. In addition to the functional importance of this complex, its structural and mechanistic complexity has further stimulated attention.

Fig. 1. Role of the F_1F_0 ATPase complex in membrane energetics of facultative bacteria.



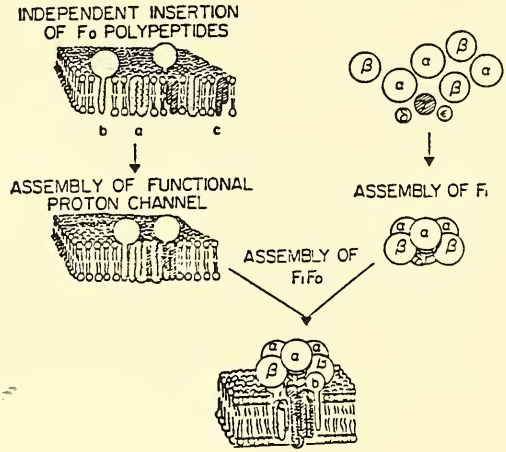
NIGMS grantee, Dr. Robert D. Simoni, has had a distinguished career studying the structure and function of the Escherichia coli F_1F_0 ATPase. As diagrammed in Figure 2, the enzyme consists of two components. F_1 is peripheral to the membrane, and easily removed and purified. It is composed of five different polypeptides, which are present in different stoichiometries (3:3:1:1:1). The F_0 portion is an integral membrane component with 3 protein constituents present in molar ratio of 1:2:10.

Fig. 2. Subunit molecular weight and stoichiometry of the Ecoli F_1F_0 ATPase.



Dr. Simoni's group was one of several which identified, cloned, and sequenced the gene for each of the eight proteins in the ATPase complex. All of these genes were tandemly clustered in a single operon termed unc (for uncoupled). Dr. Simoni proposed that the stoichiometric relationships of the various proteins in the mature enzyme were reflected in their relative rates of gene translation. He demonstrated the presence of regions of secondary structure within the appropriate intergenic spaces of the unc mRNA restrict translational initiation. Sight-directed in vitro mutagenesis, coupled with analysis of in vivo translation, supported his hypothesis. Having shown that the F_1F_0 polypeptides are synthesized in the correct relative numbers, Dr. Simoni next addressed the question of how they are assembled into a functional membrane complex. He used chromosomeless mini-cells containing plasmids carrying varying portions of the unc operon and then did cell fractionation and enzyme function studies to assess the degree of assembly. Figure 3 summarizes the results of a very large body of work:

Fig. 3. Sequence of assembly of the E. coli F_1F_0 complex.



The F_1 and F_0 portions could assemble independently. The F_0 polypeptides would spontaneously insert into the membrane and form an active, open proton channel. F_1 could also assemble independently, to form an active ATPase, but was incapable of membrane association in the absence of complete F_0. In vitro, this cytoplasmic F_1 was capable of binding to a normal F_0. Dr. Simoni systematically deleted each of theF_1 substituents to assess their role in assembly and membrane attachment. In the absence of either the δ or ϵ polypeptides, a catalytically active F_1 could form, but it was incapapble of binding to F_0. These in vivo experiments were confirmed by in vitro reconstitution studies, supporting the simple proposal of independent assembly outlined above.

Having found that F_0 forms the actual proton channel, Dr. Simoni now plans to turn his attention to unravelling the mechanism of proton translocation. His approach will be to analyze the effects of mutations which directly abolish translocation (and do not affect subunit interactions or assembly). Brute force random mutagenesis of the gene encoding the b subunit did not produce mutants with the desired phenotype. But Dr. Simoni has been able to identify two separate amino acid substitutions in the a polypeptide which abolish proton pumping. With this encouraging observation he will now be able to initiate site-directed mutagenesis to further explore the topology and function of F_0.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date: December 28, 1988

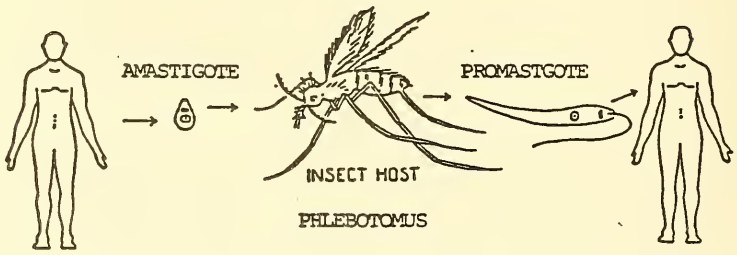
From: Program Administrator, CMBD, NIGMS

Subject: Research Highlight: "The Proton Pump of Leishmania donovani: Life in the Acidic Vesicle" R01 GM39622 (Stringer, J.R.), University of Cincinnati

To: Director, NIGMS
Through: Director, CMBD, NIGMS

Leishmania donovani is a parasitic protozoan which afflicts over 11 million people worldwide. It is the causative agent of visceral leishmaniasis (or kala azar), a fulminating, rapidly metastasizing infection of reticuloendothelial cells of the liver and spleen. It is generally fatal. Current therapy is hampered by the toxic side effects manifest by antileishmanial drugs and the development of resistant strains.

L. donovani has a complex life cycle involving two distinct morphological forms (see Figure 1). A flagellated extracellular form, the promasigote, lives in the intestine of the invertebrate host (sandflies of the genus Phlebotomus). When an infected fly takes a blood meal, it regurgitates the promastigotes into the host's bloodstream. The circulating promastigotes are rapidly phagocytized by host macrophages, and other reticuloendothelial cells, where they transform into a small non-flagellated variant called the amastigote. Intracellular growth and multiplication of amastigotes results in host cell destruction and is the basis of the pathobiology of the disease.



Host reticuloendothelial cells attempt to destroy Leishmania amastigotes by taking the organisms up into phagolysosomal vesicles. Unfortunately, the parasite is able to grow and multiply in this acidic (pH 1-3) vacuolar environment which is inimicable to other microorganisms. NIGMS grantee, Dr. James R. Stringer, has spent several years investigating the molecular mechanisms which allow Leishmania to adapt to an acidic in vivo environment. In characterizing the physiology of amastigote growth (in a Syrian golden hamster model system which he developed in collaboration with Dr. Theresa A. Glaser) it was discovered that nutrient uptake and metabolism are optimal at acidic pH. This hydrogen ion effect was manifested at the level of membrane transport. Dr. Stringer found that amastigote active transport of metabolic substrates is driven by a proton electrochemical gradient (or proton motive force) across the parasite membrane. This observation gave new significance to the extremes of pH encountered by the parasite. Dr. Stringer reasoned that the highly acidic lysosomal environment should be viewed as being exploited to make a contribution to the force required to drive the uphill transport of nutrients into the parasite. And he postulated that the effect of protons on metabolic activity would probably be due to proton-coupled synporters. Therefore, he proposed a molecular characterization of L. donovani proton pumps.

In his first experiments, Dr. Stringer developed procedures to purify parasite membranes and demonstrated the presence of a single cation-transporting ATPase. Since its biochemical properties were similar to those of the well-characterized Escherichia coli aspartyl phosphate class of cation-transporting ATPases, Dr. Stringer used an oligonucleotide probe from the bacterial gene to clone the parasite enzyme. The predicted L. donovani protein contained all of the domains expected to be present in an aspartyl phosphate ATPase, with the appropriate number and distribution of hydrophobic regions. Evolutionary comparison with the fungal and mammalian enzyme classes demonstrated that the leishmanial protein has the appearance of a chimera, with patches of sequences from both lower and higher eukaryotic ATPases. Inclusion of the L. donovani ATPase in theses evolutionary lines lends support to previous conclusions about phylogenetically conserved domains in this class of enzyme, and suggests the existence of additional conserved domains not previously apparent. This is a striking example of how addition of an intermediate form can serve to link regions of proteins from distantly related organisms.

Southern blots of genomic DNA revealed that the L. donovani ATPase gene is organized as a tandemly repeated gene pair, designated as ATPase 1a and 1b. Sequencing demonstrated that the two copies were nearly identical, differing at only 20 of 974 amino acid residues. Fifteen of these alterations were clustered at the COOH-terminal. The two gene copies could be distinguished because they give rise to RNA transcripts which differ in size (5.2 vs. 5.8 kb). Northern blot analysis showed that although the 1a gene was expressed in both promastigotes and amastigotes, the 1b gene is predominantly expressed in the intracellular amasigotes. This finding suggests that the proton pump is important in adaptation of amastigotes to life in the acidic environment, and

that increased gene dosage is one mechanism for regulating protein production. Further analysis of these two genes might reveal the underlying mechanisms by which stage-specific gene expression in *Leishmania* is controlled. Dr. Stringer proposes next to overproduce both the 1a and 1b enzymes in order to determine the biochemical consequences of the structural differences between the two isoforms.

Like all parasites, *L. donovani* is endowed with the means to evade host defense systems. Residence in the lysosomes of the very cells of these defense systems is a particularly impressive selective strategy. But Dr. Stringer's work has provided much of the knowledge required to understand the molecular basis of this particularly virulent parasitism and has identified a rational target for subsequent development of chemotherapeutic agents.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date May 26, 1989

From Program Administrator, CMBD Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Acidification and the Role of Vacuolar H+-ATPase in Protein Sorting in the Yeast Cell." ROI GM 38006 (Stevens, T.H.), University of Oregon

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Proteins destined for the extracellular environment, plasma membrane, or lysosome of eukaryotic cells are directed to their ultimate site of function by a series of membrane-enclosed organelles that comprise the secretory pathway. The set of reactions which lead to intercompartmental transfer, sorting to various destinations, and covalent modification of such proteins is complex, and clearly depends on the action of large numbers of cellular gene products. The simple yeast, Saccharomyces cerevisiae, has emerged as a useful genetic system for investigating the mechanism of lysosome biogenesis. The yeast cell has an acidic vacuolar organelle, containing numerous hydrolytic enzymes, which is analogous to the lysosome of higher eukaryotes. NIGMS grantee Dr. Tom H. Stevens, University of Oregon, has been one of many cell biologists working to characterize the molecular details of vacuolar protein trafficking.

Dr. Steven's laboratory began by identifying the genes encoding three of the best characterized soluble vacuolar enzymes: carboxypeptidase Y (CPY), proteinase A (PrA), and proteinase B. From the deduced sequence of these genes it was clear that, like their lysosomal counterparts, all three enzymes were synthesized as glycosylated, larger molecular weight, inactive precursors. These proenzymes undergo post-translational modifications duringtransport through the Golgi apparatus and then are proteolyically cleaved to their mature forms upon vesicular transport to the vacuole. Using deletion anaylsis and site-directed mutagenesis, Dr. Stevens identified a 12 bp NH_2 - terminal topogenic element, immediately following the endoplasmic reticulum (ER) signal sequence, which is necessary for targeting CPY and PrA to the yeast vacole. As has been the case with other localization sequence systems (ER, nuclear, peroxisomal) these two vacuolar determinants shared no primary sequence similarity.

In order to determine what as yet unrecognized structural features might be shared by the vacuolar topogenic elements, Dr. Stevens began a study to identify the cellular constituents of the vacuolar sorting machinery which must interact with the targeting signals. He has employed two distinct methods to discover such components of the sorting apparatus:

1. Mutational Disruption

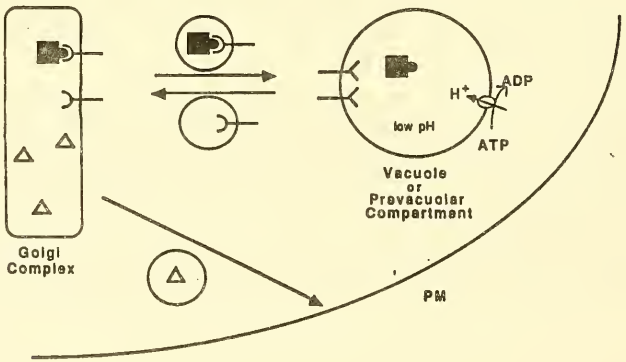
Dr. Stevens developed a genetic selection for cells which mis-localize CPY to the cell surface, rather than the vacuole. He isolated a vast number of mutants, constituting a minimum of 49 distinct complementation groups, called

vacuolar protein sorting (VPS), whose products are required to sort, transport, and/or retain vacuolar proteins. Obviously, this is an extremely complex apparatus. Determining the biochemical role of each specific VPS gene product will be a major undertaking. Dr. Stevens has begun by cloning several of these genes and by developing an in vitro assay for Golgi-to-vacuole protein transport to screen their activity. As a guide to selecting particular VPS genes for in-depth study, Dr. Stevens attempted a second procedure for impairing vacuolar localization.

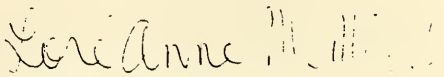
2. Physiological Disruption

Since the vesicular organelles known to participate in protein transport and processing reactions maintain an acidic interior, Dr. Stevens analyzed the effects of two agents that abolish acidification of the vacuolar lumen: ammonium chloride, a classic lysosomotropic agent, and bafilomycin A_1, a specific and potent inhibitor of the vascuolar H+-ATPases which generate and maintain the acidic pH. He found that both agents caused mislocalization of CPY to the cell surface. Dr. Stevens next screened the complete set of VPS mutants to determine their effect on vacuolar pH. Two of the 49 classes, VPS3 and VPS6 were required for the establishment or maintenance of a low vacuolar pH. Moreover, the vacuolar membranes from both mutant cells were deficient in at least two of the subunits of the vacuolar H+-ATPase (although the whole cell steady-state levels were equivalent to wild type.) Therefore, some of the VPS gene products are required for the normal function of the ATPase complex. They may either encode enzyme subunits or components required for their assembly and targeting.

While this genetic analysis in yeast has revealed that the vacuolar H+ - ATPase is required for vacuolar protein sorting, it has not elucidated the precise nature of its role. Dr. Stevens has proposed the following model.



Failure to acidify the vacuole or a prevacuolar (uncoupling) compartment results in recycling of receptors back to the Golgi apparatus in a ligand-bound state. Newly synthesized vacuolar proteins would thus encounter only bound receptor and be secreted by the bulk-flow secretion pathway by default. Although many aspects of this model remain to be tested, the tools are available to begin a critical dissection of its predictions.



Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 21, 1989

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlight: "The Yeast Osmosensor, a Stretch-activated Ion Channel." *
ROI GM37925 (Kung, Ching), University of Wisconsin

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

One of the major functions of the cell membrane is to provide a permeability barrier which maintains the appropriate intracellular ionic environment. The membrane can "sense" external changes in osmotic pressure, and transmit this information to regulatory molecules which orchestrate compensatory responses. Past highlights in this series have described several types of alternate osmoregulatory systems in microorganisms. These systems induce transport and/or biosynthetic pathways which rapidly alter the colligative properties of the cytoplasmic and periplasmic fluids in order to adapt to environmental disruptions in osmolarity. Despite these considerable advances in our understanding of osmoregulatory loops, the molecular nature of the osmosensing apparatus has remained elusive. Indeed it has been conceptually difficult to posit a unitary "reporter" molecule(s) whose presence or concentration would uniquely measure osmotic pressure and which could activate a surface receptor osmosensor.

NIGMS grantee Dr. Ching Kung, University of Wisconsin, has recently made an exciting discovery which provides the first solid description of an osmosensor. He has found that both the yeast, Saccharomyces cereviseae, and the bacterium, Escherichia coli, contain mechanosensitive, or stretch-activated, ion channels which respond directly to turgor pressure. Dr. Kung's achievement has come with a long and indirect intellectual history, illustrating how unexpectedly basic research in one field can lead to major advances in another.

Dr. Kung's life long work has centered on a biochemical genetic study of behavior at the cellular level. Since organismic behavioral responses are governed by electrical currents, he chose the minimalistic model system to which patch clamp techniques could be applied: the ciliated protozoan Paramecium. This giant beast combines extraordinary beauty with all the oft-cited benefits of microbial systems (fast growth, ease of cultivation, accessible genetics). Dr. Kung spent many years isolating series of behavioral mutants ("dancer", "fast", "pawn", "paranoic") and characterizing their biochemical defects, which turned out to reside in various types of ion channels. Ion channels are a special class of integral membrane protein which act as switches. When open, channels allow the passage of ions or electrical

* The third in a series on osmoregulation from this program.

current. Since an open channel will dissipate the electrochemical gradient established by membrane pumps, the ion channels are "gated", and the gates are usually closed. The probability of opening the gate can be modified by specific signals. More than 60 different types of ion channels are currently recognized, and they are classified by the types of signal which gate them (external or internal ligands, other membrane proteins, voltage). In higher organisms, ion channels underlie smell, taste, hearing, and vision, as well as responses to hormones and neurotransmitters and the function of smooth muscle networks.

Dr. Kung described eight different types of ion channels which affect the behavior of Paramecium. In the course of this work, Dr. Kung refined his microelectrode technique to the point that he was able to develop procedures to perform patch clamp studies on smaller cells of the yeast S. cerevisiae. He removed the cell wall with the enzyme zymolase, yielding spheroplasts which could form gigaohm seals if sustained, vigorous suction were delivered to the patch clamp pipet. Once the seal was formed, Dr. Kung could measure two types of prominent ion channels in the yeast cell. One was a typical K+, voltage-sensitive channel. But the second discharged picoampere currents in response to fluctuations in the suction on the pipet. This channel responded to pressures equivalent to a few centimeters of mercury; the stronger the suction, the more likely the channel was to be open. Dr. Kung quickly demonstrated that this pressure-sensitive channel could also be activated by osmotic pressure. Dilution of the solution bathing the spheroplast (creating an outward osmotic pressure) activated the channel. He showed that these channels preferred cations over anions, and were not particularly selective towards various cations. The average yeast cell contains approximately 100 of these mechanosensitive channels.

Dr. Kung postulated that the physiologic role of the mechanosensitive channel was to function as an osmosensor; but he recognized that the ideal model system to test this hypothesis would be the bacterium E. coli, where the genetics and biochemistry of osmoregulation are well developed. Unfortunately, bacterial cells are far too small for patch clamp experiments. So Dr. Kung collaborated with his neighbor, Dr. Julius Adler (AI08746), to produce giant E. coli spheroplasts from penicillin-induced filaments. Such spheroplasts were 5-10 μm in diameter and Dr. Kung could achieve patch-clamp recordings. He found that E. coli also contained a pressure-sensitive ion channel. Interestingly, the bacterial channel was voltage-sensitive, and pressure exerted in either direction could open the channel.

The evolutionary origin of mechanically-gated channels (which in vertebrates are responsible for hearing) has now been traced to microbes. It is possible that the detection of water through osmotic pressure was the first "sense" of all life forms. Dr. Kung and his collaborators are now embarked on studies to purify, sequence, and determine the function of the yeast and E. coli mechanosensitive channels. This work should continue to expand our molecular understanding of both behavior and osmoregulation.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 25, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlight: "The Membrane Transport of Vitamin B12."
GM19078 (Kadner, Robert J.) University of Virginia, Charlottesville, Virginia.

To Director, NIGMS
Through: Director, CMBD Program, NIGMS BIS (8/25/88) for CAM

The structure and function of cobalamin (Vitamin B_{12}) have intrigued students of human biology since 1926, when Minot and Murphy demonstrated that pernicious anemia could be controlled by eating one-half pound of raw liver per day. This fatal syndrome, largely a disease of strict vegetarians or the elderly, was associated with a reduction in the number of erythrocytes and other rapidly growing tissues. Though effective, the treatment was one which obviously not all patients accepted with enthusiasm. And this spurred the pharmaceutical industry to isolate and characterize the active factor. By 1956, the study of cobalamin chemistry culminated in Dorothy Hodgkin's elucidation of its three dimensional structure. As illustrated in Figure 1, the vitamin is composed of a central cobalt atom surrounded by a planar corrin ring with a complex side chain below. The molecule is completed by coordinate linkage of any one of several radicals to the cobalt nucleus.

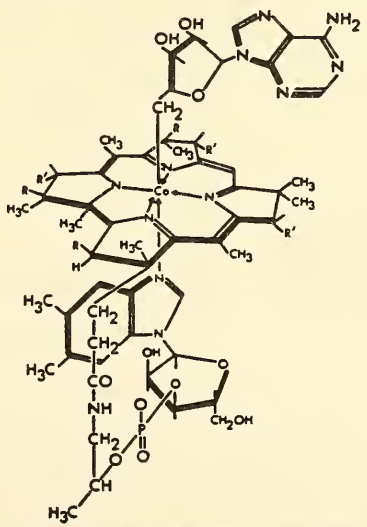


Figure 1 Structure of adenosylcobalamin (AdoCbl). R = CH_2CONH_2; R' = $CH_2CH_2CONH_2$. Other radicals which may be coordinately linked to the cobalt atom include CH_3 (methylcobalamin), OH^- (hydroxocobalamin), and CN^- (cyanocobalamin).

Figure 1

Although this unusual nutrient is widely distributed in animal tissues, it is not synthesized by eukaryotic cells. Rather, it is produced and excreted by intestinal microorganisms. Pernicious anemia is not always the result of a dietary deficiency of cobalamin but is frequently associated with failure of the patient to absorb the vitamin from intestinal fluid. This absorption depends on the combined action of gastric, ileal, and pancreatic components. First, cobalamins are released from dietary protein in the acid environment of the stomach, where they bind to "R" glycoproteins from the saliva. Pancreatic proteases next digest the R proteins to liberate cobalamins in the ileum. Here, they complex with "intrinsic factor" (IF), a family of mucoproteins secreted by gastric parietal cells. The IF-B_{12} complex interacts with a surface receptor in the presence of Ca++. In this process, the complex is dissociated and B_{12} crosses the ileal membrane to the portal blood. Once in the bloodstream, the free vitamin binds to one of another set of R proteins, the three β-globulin transcobalamins or hepatocobalamin. These circulating complexes are recognized by specific plasma membrane receptors and internalized by endocytosis. As outlined in Figure 2, lysosomal proteases free the B_{12} either to enter the mitochondrion and participate as a coenzyme for methylmalonyl CoA mutase (an enzyme vital for the catabolism of fat and protein) or to remain in the cytoplasm as a cofactor for CBL-methyltransferase (a key enzyme in sulfur amino acid metabolism).

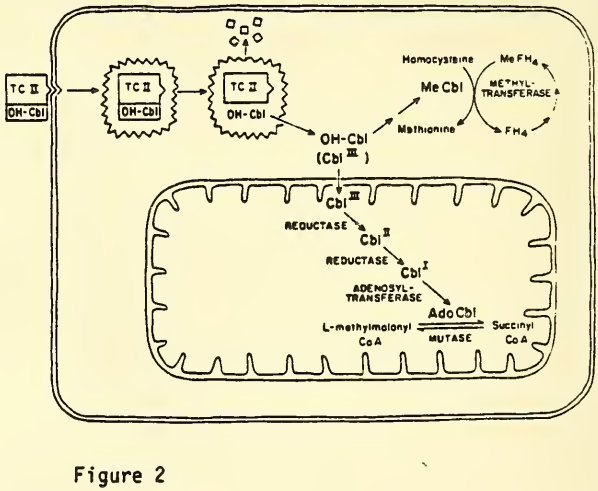


Figure 2 General pathway of the cellular uptake and subcellular compartmentation of cobalamins, and of the intracellular distribution and enzymatic synthesis of cobalamin coenzymes. Abbreviations: TC II, transcobalamin II; OH-Cbl, hydroxocobalamin; MeCbl, methylcobalamin, AdoCbl, adenosylcobalamin; Cbl^{III}, Cbl^{II}, Cbl^{I}, cobalamins with cobalt valence of 3+, 2+, and 1+, respectively

Figure 2

Obviously, the cobalamin vitamins have a unique and highly specialized mechanism of transport, which represents a site of significant disease. B_{12} itself is not a substrate for transport and is only taken up in a complex with some specific binding protein. The eukaryotic cell surface receptors for these complexes have been isolated from various tissues. They appear to be large, hetero-oligomeric structures with one subunit that is very similar to the B_{12} transport protein. This observation led to the suggestion that this

receptor subunit and its substrate are products of a gene duplication event. Unfortunately, the molecular characterization of these receptors has not yet been possible. Indeed, the structure of a cobalamin-binding domain is not known for any protein, even though it seems likely that the binding site for such a complex and essential ligand would be at least partially conserved during evolution. Since the mammalian cell system has proved so difficult to delineate, NIGMS grantee Dr. Robert Kadner (GM19078), University of Virginia, has turned to the simple bacterium Escherichia coli as a model system to investigate the molecular mechanisms of B_{12} absorption and transport.

Like the eukaryotic cell, E. coli cannot synthesize B_{12} and must transport this essential coenzyme from the environment. Dr. Kadner discovered that the bacterium employs a high-affinity, irreversible transport process, rather than relying on diffusion of this bulky compound through a porin. Dr. Kadner has used two approaches to study the details of E. coli B_{12} transport: biochemical isolation of B_{12}-binding proteins, and molecular genetics employing lacZ fusion mutagenesis to identify genes whose products are required for B_{12} uptake. His work has culminated in revealing the existence of a novel multi-step uptake process, each step involving independent transport components (as diagrammed in Figure 3):

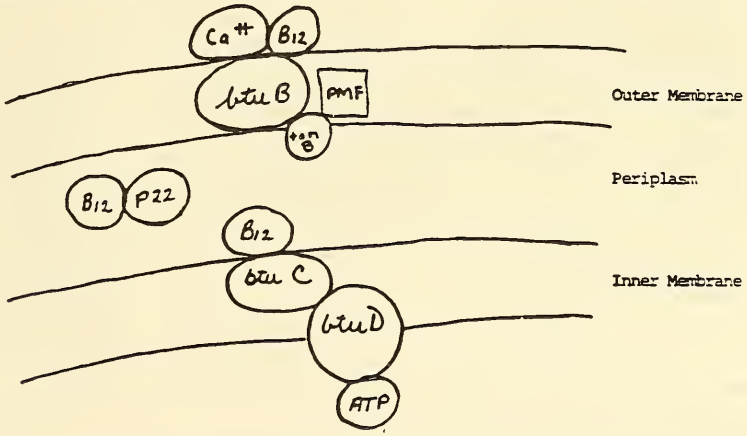


Figure 3. The Escherichia coli cobalamin transport components.

Figure 3

Outer Membrane Transport: The outer membrane vitamin B_{12} receptor is the product of the btuB gene, which Dr. Kadner has cloned and sequenced. It has a molecular weight of 66,400 and requires Ca++. The btuB protein carries a single binding site which acts as the receptor for such lethal entities as phage BF23, the E colicins, and colicin A, as well as cobalamin. It is a minor membrane component, present at a level of 200-300 molecules per cell. The expression of btuB is under the control of a regulatory locus btuR, which encodes a 21kd protein. Dr. Kadner has also cloned and sequenced this gene

and described a B_{12}-dependent negative repression loop. Release of B_{12} from the btuB protein into the periplasm is a proton motive force (pmf)-dependent process which requires the presence of the tonB gene product (although the use of the btuB protein by the phage and colicins does not require tonB). Dr. Kadner has cloned and sequenced the tonB gene. It not only plays an essential energy-dependent transducing role in B_{12} uptake but also functions in the transport of various iron-siderophore complexes by various ferric chelate outer membrane receptors (fepA, fhuA, and iutA).

Periplasm: There is a 22kd periplasmic B_{12}-binding protein which can be released from E. coli cells by osmotic shock. The genetic locus which encodes this protein has not yet been identified, but Dr. Kadner has purified the protein.

Inner Membrane: The btuC gene encodes a 31,683 Mr integral inner membrane protein with structural homology to other E. coli periplasmic substrate-binding protein-dependent transport systems (histidine, maltose, phosphate, and oligopeptide). The btuD gene encodes a 27,088 Mr polypeptide with extensive homology (46%) to the ATP-binding peripheral membrane proteins associtated with these four systems (hisP, malK, pstB, and oppD).

To further characterize these components of the B_{12} transport complex, Dr. Kadner has begun in vitro mutagenesis studies to define segments of the various proteins which control the binding of individual ligands. In the btuB gene, he has identified a "TonB box", eight amino acids near the amino terminus which are required for btuB interaction with tonB. He located the same box sequence in three other tonB-dependent receptor genes, fepA, fhuA, and iutA. Dr. Kadner also has identified the btuB residue (thr50) responsible for Ca++ binding. In the btuD gene, Dr. Kadner has identified sequences which affect ATP binding. Two regions of btuD show 50% sequence homology to btuB (reminiscent of the duplicate homologies in the eukaryotic system), and provide candidates for future exploration of B_{12}-binding domains. Obviously his laboratory is now poised to complete the molecular dissection of the complex B_{12} uptake system.

Lore Anne McNicol

Lore'Anne McNicol, Ph.D.

gb

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Deputy Director, CMBD, NIGMS

Subject Research Bullet: "Genetic Study of Ion Transport in Neurospora"; GM 15761-20 (Carolyn W. Slayman) Yale University, New Haven

To Director, NIGMS
Through: Director, CMBD, NIGMS

There are several extremely important membrane transport ATPases in mammals. These include the Na/K-ATPase of nerve, muscle and heart plasma membranes, and the Ca-ATPase of sarcoplasmic reticulum. The former is responsible for maintaining the membrane potential in nerve and muscle, and is the target of some antiarrhythmia drugs, such as digoxin. The Ca-ATPase plays a critical role in recovery after contraction of skeletal or cardiac muscle. Some recent evidence implicates decrease of this activity in the loss of cardiac output with aging. Another member of this family is the plasma membrane ATPase of the mold, Neurospora. It has been shown to share many features with the other ATPases, namely approximate molecular weight, amino acid homologies, membrane orientation, and certain kinetic features. It has several advantages over the others, however, for detailed study. The gene has already been cloned. A yeast expression system is being worked out currently. Advances in the study of yeast genetics and physiology are allowing the isolation of suitable yeast membrane vesicles containing function Neurospora ATPase.

Recent work, using a variety of group-specific reagents has allowed workers to pinpoint individual amino acid residues that participate in the reaction cycle. For example, it is now known that glutamate-129, which is located in the middle of the first transmembrane segment of the enzyme, is essential to the reaction cycle, probably transporting a proton. Lysine-474, in a site found in all the homologous membrane ATPases (including mammalian) is essential, and close to the site at which ATP is known to phosphorylate the enzyme (aspartate-378). In short, we are well on our way, using recombinant DNA technology coupled with sophisticated biochemistry and physiology, to understanding the workings of an interesting membrane transporter and enzyme which is an excellent model for essential and less well understood mammalian tranporters.

Bert I. Shapiro, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 4, 1988

From Program Administrator, CMBD, NIGMS
Through: Program Director, CMBD, NIGMS

Subject RESEARCH BULLET: Glycolipid Membrane Anchors - A New Class of Surface Molecules (Silver, Jack, R01 GM38207-02, Hospital for Joint Diseases, New York; Low, Martin, R01 GM35873-04, Oklahoma Medical Research Foundation; Flaherty, Lorraine, R01 GM37947-02, NY State Department of Health; Thompson, Linda, R01 GM39699-01, Scripps Clinic, California)

To Director, NIGMS

Traditional peptide receptors generally consist of an extracellular hydrophilic domain, a hydrophobic membrane spanning portion, and a more hydrophilic intracytoplasmic tail. The cytoplasmic tail has been considered a key part of the molecule for transmitting signals from the receptor-ligand bound complex to effector molecules in the cell, such as G proteins and second messengers. However a new class of molecules has recently been described which seems to be associated with cell activation events, yet has no intracytoplasmic tail. Rather, they are glycoproteins inserted into the plasma membrane via a glycolipid anchor consisting of a phosphoinositide-glycan linkage. The peptide is covalently linked to the glycolipid moiety via its carboxyl terminus. At least some of these surface activation molecules are expressed on cells of the immune system, such as Thy-1, Ly-6, and the Qa antigens; others are expressed in nervous tissue and on the surface of micro-organisms. Thus, a growing class of membrane glycoproteins, some of which are likely to play a role in cell adhesion and intercellular communication, also have been shown to be glycolipid anchored. This family of molecules includes the variable surface glycoproteins (VSG) of trypanosomes, decay accelerating factor (DAF), neural cell adhesion molecule (NCAM), the scrapie prion protein, and a series of biologically important enzymes (such as acetyl cholinesterase, alkaline phosphatase, ornithine decarboxylase, and NADH reductase. Obviously this is a diverse group of molecules whose primary structures show little homology.

In the case of VSG and DAF, and perhaps some other family members, the enzyme phospholipase D can cleave the glycoprotein moiety from the membrane anchor. The exciting questions which these findings raise, center on the mechanism whereby these molecules transmit signals. One possibility of course is that the surface glycoprotein once solubilized by cleavage from its anchor, will associate with another receptor either on the same cell or a different one.

Another possibility is that the remaining glycolipid anchor will translocate and generate second messenger molecules such as PIP3. In summary, a new class of surface molecules with a unique mode of membrane anchorage has been described. The characteristics of these glycolipid anchored molecules suggest a new pathway for transmission of signals in a wide array of cells including those of the nervous and immune systems.

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 4, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlights: "Gap FRAP: A New Technology to Study Cell-Cell Communication"
ROI GM30158 (Schindler, M.), Michigan State University

To Director, NIGMS
Through: Director, CMBD, NIGMS

Gap junctions are low resistance pathways for intercellular communication, formed after cell-cell contact. These transport channels can exchange low molecular weight substances (less than 1700 daltons); and it has been suggested that they play a significant role in the control of cell proliferation, tissue metabolism, and synchrony. The study of gap junction transfer of signals has involved the use of a wide variety of investigative methods, including electrocoupling, dye transfer, metabolic cooperation (cross-feeding), and autoradiography. In most instances a microelectrode or pipette is introduced into the cell to deliver the measurable signaling agent. The possibility for cell trauma, the difficulty of extending these technologies to all types of cells and cell aggregates, and the requirement for microinjection expertise have limited the usefulness of most analytical techniques to measure gap junction transfer.

Dr. Melvin Schindler, Michigan State University, has developed a new method to measure gap junction transfer that overcomes these problems and provides a means of flux quantitation. The procedure takes advantage of the known transfer of fluorescein dyes through the gap junction, and extends the technique of fluorescence redistribution after photobleaching (or FRAP) to whole cells. In this process, termed "gap FRAP", cells in tissue culture are labeled with 6-carboxyfluorescein diacetate. All cells in the population are internally labeled by this stain, which is maintained in the cell. Any labeled cell may then be photobleached by a laser beam whose width is approximately equal to the cell diameter or by a series of laser pulses each with a diameter of 1 um. After photobleaching, the bleached dye molecules from one cell and the unbleached dye molecules from an adherent contacting cell may be redistributed through gap junctions. Monitoring the redistribution of these labeled reporter molecules as a function of time results in a single exponential recovery curve that yields a rate constant for dye transport.

Dr. Schindler and his colleagues demonstrated the utility of gap FRAP procedures in a study of the effect of tumor promoters on cell-cell communication. Human teratocarcinoma cells in tissue culture were treated with the fluorescein dye, and subsequent fluorescence microscopy revealed homogenous cytoplasmic staining. Dr. Schindler then used laser bleaching to destroy the dye in a single contacting cell and in an isolated, unconnecting cell. Destruction of dye fluorescence emission was observed in both bleached cells. However, only the contacting cell, with functional gap junctions, recovered fluorescence. This experiment was then repeated in the presence of the pesticide dieldrin and the tumor promoter 12-O-tetradecanoyl phorbal-13-acetate (or TPA)--agents reported to inhibit gap junction communication. Under these conditions the contacting cell did not regain fluorescence.

The dye and labeling conditions for gap FRAP do not affect cell viability, and the same cells can be restained for several days. Gap FRAP permits multiple measurements of the same cell without traumatic manipulations. Perhaps the full potential of gap FRAP rests in its ability to make possible measurements of all cell types and cell configurations. The possibility of continuously monitoring communication patterns in the same developing blastomere, a measurement that would appear to be enormously difficult with existing methods, may now be considered.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

ADMINISTRATIVE NOTE: The development of gap FRAP instrumentation was done in conjunction with an SBIR grant, R43 CA39945, to Meridian Instruments. This represents a very successful example of government-academe-industry collaboration in biotechnology development.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlights: "Diffusion Properties of the Gap Junction"
GM24905 (Brink, P.R.), SUNY at Stony Brook

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Intercellular communication in the form of solute transfer or ionic coupling has been found in association with the junctional membrane structure known as the gap junction in almost every tissue and cell type. A number of factors are known to affect the movement of substances through the gap junction: including voltage, phosphorylation, $[H^+]$, $[CA^{++}]$, and various pharmacologic agents which alter cellular ionic concentrations. All of these factors are thought to act through sensors or "gates", which presumably are an integral part of the gap junction channel. The structure of the isolated channel has been partially determined by x-ray analysis. But the lack of the x-ray opaque tracer molecules that can penetrate the channel has prevented a detailed look at channel ultrastructure. Biochemical analysis has produced the partial sequence of junctional membrane protein subunit, and this is consistent with a hydrophilic intercellular channel. Diffusion of hydrophilic probes through the gap junction, as well as high ionic conductance, also lend support to the hydrophilic channel concept.

Dr. Peter R. Brink, SUNY at Stony Brook, has contributed significant evidence supporting the hydrophilic gap junction model. In the past, his laboratory determined that the permeance of junctional membranes to fluorescent probes was consistent with passive, hydrated diffusion through a solvent-like environment. His observation suggested that the cytosol has a role in regulating junctional membrane permeability. Recently, he has focused more narrowly on the mechanisms of junctional diffusion.

Since passive diffusion processes are largely determined by solvent viscosity, Dr. Brink embarked on a series of experiments to measure solvent effects on gap junction communication. His model system is the earthworm septate axon, which is composed of large cylindrical cells that are joined end-to-end with extensive areas of gap junctions. The axon is extremely well-suited to microinjection studies. Dr. Brink examined the effects of substituting the natural biological solvent (H_2O) with deuterium oxide (D_2O), which increases the viscosity of the cytosol. D_2O rapidly reduced both the axoplasmic and gap junctional conductances, with similar initial reduction rate constants. Analogous experiments were done with temperature, glycerl, and polyethylene glycol--agents which also increase solvent viscosity. Taken together, these experiments consistently demonstrated that ion mobility through the channel is primarily governed by viscous properties of the solvent. Solvent clearly occupies the gap junction channel volume and influences transjunctional ion mobility. Dr. Brink's findings are inconsistent with models of a non-aqueous gap junction channel or a transport mechanism that occurs via ion binding. Thus, he has provided an important new look at the physiology of gap junction channels.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

Memorandum

Date March 4, 1988

From Program Administrator, CMBD, NIGMS
Through: Program Director, CMBD, NIGMS

Subject RESEARCH BULLET: Intercellular and Intracellular Signal Transduction in Yeast (Herskovitz, Ira, P01 GM31286-06, University of California at San Francisco; Reed, Steven, R01 GM38328-02, Scripps Clinic & Research Foundation, La Jolla, California; Sprague, George, R01 GM38157-03, University of Oregon, Oregon; Thorner, Jeremy, R01 GM21841-14, University of California at Berkeley)

To Director, NIGMS

Yeasts reproduce either by a sexual or an asexual mating process, This simplest of eukaryotes has two sexual mating types, a and alpha, each of which produces a pheromone with which it signals to its reciprocal mating type. These pheromones, termed a and alpha-mating factor, are small peptides which bind to specific receptors on the opposite cell type, thereby causing a growth arrest and further development. The structure of the peptide pheromones is now known and the nature of the receptor to which these peptides bind is a subject of great interest. Studies actively supported within NIGMS address all steps in the signal transduction pathway leading to growth arrest, from regulation of receptor expression, to receptor-ligand binding sites, to signal transfer via putative G proteins, to the second messengers which transmit the signal to the nucleus, to the activation and repression of gene expression, which ultimately leads to differentiation.

Since the gene for the alpha-factor receptor (STE-2) has been cloned and sequenced, the primary amino acid sequence of the receptor molecule is known and has been shown to contain alternating hydrophilic and hydrophobic stretches which indicate that the receptor is a seven membrane spanning receptor similar to those described for the rhodopsin and B adrenergic receptors of higher organisms. Dr. Thorner and Dr. Sprague are trying to identify the functions of each of the key domains of the receptor molecule by creating mutant genes or peptide fragments and either expressing the altered gene products in yeast, or preparing antibodies to the receptor domains. Of particular interest are those domains which bind the ligand (extracellular) and those domains which bind the effector molecules (intracellular). Dr. Reed is actively searching for the yeast G proteins analagous to those which associate with vertebrate adrenergic receptors, and Drs. Herscowitz and Reed are isolating mutant mating types which are defective in their second messenger signalling pathways. To help dissect the biochemical steps which occur between the membrane and the nucleus, all of these investigators will seek ultimately to be able to understand at the gene level how these intercellular mating factors control cell growth and differentiation. In

summary the mating factors of yeast provide an ideal model system for dissecting sensory signal transduction pathways which resemble those which operate in more complex higher organisms.

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Program Administrator, CMBD, NIGMS

Subject Research Bullet: "Intercellular and Intracellular Signalling in the Slime Mold"; GM 33136-06, (Richard Kessin) Columbia University, New York.

To

Director, NIGMS
Through: Director, CMBD, NIGMS

The cellular slime mold is a fascinating organism which provides an ideal model for studying important basic mechanisms. The list of problems which can be addressed with Dictyostelium discoideum includes cell migration and cell aggregation, intercellular communication and intracellular signalling, and cell differentiation and morphogenesis. The slime mold exists in two forms, a unicellular ameoboid cell, and a complex multicellular structure consisting of a stalk topped by a fruiting body containing spores. It is the transition from unicellular to multicellular organism which provides the opportunity to study so many basic phenomena. The amoebae respond to a loss of their food source (bacteria) by swarming together and aggregating. This initial phase is mediated by synthesis and release of pulses of cAMP, normally found in cells only as an intracellular second messenger. At the same time the cAMP receptor, secretory apparatus, and enzymes needed for its signal transmission are synthesized. Once aggregated, a battery of genes become expressed leading to further cellular differentiation and development. In essence the slime mold has invented the strategy of coordinating the expression of its genes and morphogenesis by connecting a cell surface cAMP receptor to its sensory and motility apparatus and to the second messenger cascades that it uses for gene regulation. cAMP therefore has immediate (extracellular) effects on chemotaxis and motility in the manner of a neurotransmitter and more delayed but longer lasting effects on gene expression and development.

One of the key questions in this system is how the chemotactic waves and therefore aggregation are regulated. In order to address this question, Dr. Kessin decided to study the enzyme required for cAMP breakdown, (phosphodiesterase) and its companion inhibitor. He found first that the genes for these two glycoproteins are regulated by cAMP; phosphodiesterase is induced and the inhibitor gene is repressed by exposure to extracellular cAMP. He has succeeded in purifying these two proteins, sequencing them and cloning the genes that encode them, and can now turn to studies of their regulation and function. The phosphodiesterase, like the cAMP is secreted by the slime mold, unlike the situation in other cells where they are strictly intracellular. The secreted phosphodiesterase was found in several forms of differing molecular weights, which was due to its association with varying amounts of inhibitor in a molecular complex. The secreted phosphodiesterase, devoid of its inhibitor was identical to the membrane bound enzyme. The net

result of having both secreted and bound phosphodiesterase, as well as varying levels of the inhibitor, is a fine level of control for cellular aggregation. This is needed because too much constant cAMP results in desensitization (down regulation) of the cAMP receptor and too little results in failure to aggregate. When cAMP is too abundant it binds to the membrane phosphodiesterase where it is destroyed. When the extracellular levels fall the phosphodiesterase is jettisoned and inactivated by newly secreted inhibitor.

Having cloned the genes for the phosphodiesterase and its inhibitor, Dr. Kessin is in an excellent position to probe the control of cAMP in aggregation and its role in development. For example mutants have been created which greatly over-express phosphodiesterase. These mutants aggregate more rapidly but fail to develop normally. Ultimately the genes isolated for extracellular phosphodiesterase and its inhibitor can be engineered and inserted into cells to alter the intracellular second messenger levels as well. This should prove to be a powerful tool for exploring signal transduction pathways and the roles of key intracellular mediators of hormonal responses.

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 1, 1988

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "A New Pathway in Inter-organelle Traffic" R01 GM 38895 (William Wickner), University of California, Los Angeles, CA

To Director, NIGMS
Through: Program Director, CMBD, NIGMS

Vacuoles are cellular organelles whose primary function is storage of specialized cytoplasmic contents, and which respond to changes in physiological conditions, including cell growth and cell division. Yeast provides an ideal model for studying organelle regulation because there is usually only a single vacuole which is generally visible with the light microscope. The study of this organelle can be used to address an important question in cell biology of how patterns of cytoplasmic inheritance are determined when a cell divides. Much is known about how the nucleus segrates its genetic contents, yet relatively little is known about organelle division and traffic.

In yeast, daughter cells are generated by two processes - either sexual mating of two yeast cells to form a diploid zygote followed by cell division, or a vegetative process whereby a small bud forms from the parent cell. One of the fascinating aspects of this vegetative process is that the bud receives a small vacuole when it is only 1% of the mother cell volume; however about 50% of the bud vacuole contents are derived from the mother vacuole. In order to explain this unequal distribution of cytoplasm and vacuole contents, Dr. Wickner hypothesized that vacuole formation entails an active process of vesicular "traffic" between mother and bud vacuoles whereby a stream of small vesicles transport the vacuolar contents.

In order to test this hypothesis, Dr. Wickner and a postdoctoral fellow, Lois Weissman (F32 GM11564) studied "communication" between parental vacuoles within a zygote formed by sexual mating. This system takes advantage of a unique characteristic of the ade 2 strain of yeast, i.e. that the ade 2 vacuoles accumulate an endogenously formed fluorescent dye when deprived of the nucleotide adenine. Since dye synthesis is suppressed by adenine and is stable, mating between the ade 2 and the ADE+ yeast strains will result in one fluorescent vacuole and one unlabeled vacuole when the yeast are grown under appropriate conditions. Exchange of vacuolar contents can then be traced and quantitated by fluorescence microscopy. Using this approach it was found that dye transfer occurred in a narrow time frame (between 3.4 - 4.8 hours after mating) when vacuoles were a critical size (9.4% of the zygote volume). In addition transfer of dye to a newly formed bud of the zygote occurred 30 minutes after bud formation via a fluorescent vesicular trail (see photo of fluorescent trails between vacuoles and newly formed bud). This vesicle mediated process of vacuole exchange in the zygote is a continual process and occurs in the absence of vacuole fusion.

These findings demonstrate that organelle traffic can occur via a cell cycle-regulated pathway. This mechanism is in contrast to the behavior of other yeast organelles, such as the nuclei, which fuse after zygote formation. In mammalian cells, vesicle mediated traffic is known to occur during endocytosis and biosynthetic transport. Such traffic also has been documented in the mammalian Golgi. In contrast mammalian lysosomes have been reported to exchange contents following syncitia formation (cell fusion) but the mechanism by which this exchange occurs is as yet unknown. Such mammalian exchange of organelle contents may well follow this new model of inter-organelle traffic described in yeast. This new pathway also may be the mechanism by which cytoplasmic inheritance occurs in vegetatively and sexually reproducing eukaryotic cells.

Dr. Wickner's observation opens up many additional, interesting questions. Some of these questions are - why the transfer is tightly timed once it occurs, but is delayed for a while after cell fusion; whether the cytoskeleton needs to be rearranged for vesicle-mediated transfer to occur; how is the direction and target site of the vesicular traffic regulated? Hopefully future work will provide new answers to these important questions about regulation of inter-organelle traffic.

Marion Zatz

Marion M. Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 29, 1989

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlight: "Ras Proteins are Required for Yeast Secretion."
R01 GM35370 (Novick, P.J.), Yale University School of Medicine

To Director, NIGMS
Through: Director, CMBD Program, NIGMS Carl

The lower eukaryote Saccharomyces cerevisiae conserves many of the features of the protein secretion pathway of animal cells. In both systems secretory proteins are inserted into the lumen of the endoplasmic reticulum and then passed to the Golgi apparatus. Subsequent transport to the cell surface is accomplished by fusion of Golgi-derived secretory vesicles with the plasma membrane. Secretion in the yeast is a rapid, constitutive process which is restricted to the actively growing region of the cell surface, the bud. It has a transit time from the Golgi of approximately three minutes and does not require any external signal (required for stimulus-coupled exocytosis in higher systems). For this reason, late components (post-Golgi) of the secretory machinery are not a prominent feature of wild type cells, and their characterization has lagged compared to other studies of the phenomenon.

In order to address the molecular mechanisms of yeast protein secretion, many workers in the field have turned to a biochemical genetic approach. While in training with Dr. Randy W. Schekman, University of California, Berkeley (GM26755 and 36881), Dr. Peter J. Novick, now at Yale University, helped isolate a large series of temperature-sensitive yeast secretion (sec) mutants (see Highlight of 11/18/86). At their restrictive temperatures, these mutants block transport and exaggerate those secretory organelles which accumulate behind the block. Recently, Dr. Novick has focused his attention on the sec 4 mutant which acts late in the secretory pathway, causing accumulation of post-Golgi vesicles. He cloned the sec 4 gene from a wild type yeast plasmid library by complementation, and determined its DNA sequence. The predicted protein had a molecular weight of 23.5 kD. A database search for homologous sequences revealed a surprising and significant homology between sec 4 and all ras transforming proteins.

Ras was originally identified as a viral transforming gene responsible for neoplastic proliferation. It is now understood also to be related to normal cell constituents, part of a diverse and ubiquitous family found in eukaryotes from yeast to man. These proteins appear to function in signal transduction, involving GTP binding. Ras proteins are integral membrane constituents, where they interact with external ligands. Ras-ligand binding leads to GTP-for-GDP exchange on the ras protein. In the GTP-bound state ras undergoes a

conformational change and can stimulate a downstream effector protein. This activation is curtailed by slow hydrolysis of the bound GTP. The ras/sec 4 homology was found over the full lengths of the proteins, and was particularly pronounced in the GTP-binding and GTP-hydrolysis domains. Replacement of the chromosomal sec 4 gene with a null allele created by in vitro mutagenesis demonstrated a recessive lethal phenotype. Therefore, sec 4 is an essential cell function; and it does not play an interchangeable role with the other three known yeast ras-like homologues. Dr. Novick's physiologic studies showed that sec 4 causes a rapid block in secretion, consistent with a direct role in this process. Analysis of double mutant phenotypes provided evidence that the sec 4 protein can suppress the defects of three other late sec genes, suggesting that it may function as a master regulator.

Dr. Novick's findings raise the question of why secretion requires such a putative signal transduction protein, particularly since yeast have a constitutive pathway that requires no external stimulus. Dr. Novick is actively investigating three hypotheses to explain the observed GTP-binding domain homology between ras and sec 4:

1. The homology has no functional significance. Even though the sequence predicts GTP binding, sec 4 has diverged so significantly from ras that this site plays no role in sec 4's involvement in protein secretion.

2. Even though the yeast secretory pathway has no temporal stimulus regulation, sec 4 may function in the spatial regulation which restricts protein secretion to the bud.

3. Sec 4 may actively regulate all effector enzymes of the constitutive pathway at a constant steady state level.

Dr. Novick is presently purifying the sec 4 protein in order to test the first hypothesis through direct biochemical measurement of GTP binding and/or hydrolysis. He will also obtain antibodies to the sec 4 product to examine its cellular localization and further address its biochemical function.

Lore Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

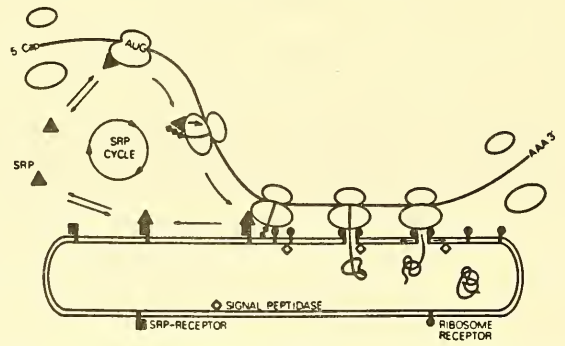
Date March 6, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlights: "The Mechanism of Protein Translocation Across Membranes"
R01 GM32384 (Walter, P.), UCSF

To Director, NIGMS
Through: Program Director, CMBD, NIGMS

An important aspect of the biosynthesis of many proteins is their intracellular topography. Many proteins have to cross or become asymmetrically integrated into membranes in order to assemble properly and reach their final functional destination. One particular group--all proteins which are excreted to the cell's exterior space--is synthesized on membrane-bound ribosomes on the rough endoplasmic reticulum (ER). Our current understanding of how ribosomes synthesizing these proteins are specifically targeted to ER is largely based on work initiated by Dr. Peter Walter while working in the laboratory of Dr. Gunter K. Blobel (GM27155), Rockefeller University. They developed techniques to reconstitute this process in vitro and to characterize biochemically the components involved. This work led to the following model for protein translocation across the ER:



A small cytoplasmic ribonucleoprotein, the signal recognition particle or SRP (▲) recognizes the signal peptides (▬) characteristic for external localization on nascent protein chains as they emerge from the ribosome. Concomitant with this signal recognition event, SRP arrests further translational elongation, retaining the nascent chain. Upon interaction of the "arrested" ribosome-SRP complex with the correct ER target membrane

containing an SRP receptor or docking protein (▬), the elongation arrest is released. A ribosome-membrane junction is established, allowing translocation of the nascent chain across the membrane. The signal peptide is removed by the membrane-associated signal peptidase (◇) and translocation is completed. To obtain a molecular understanding of how protein translocation is accomplished, it will be necessary to characterize the cellular machinery involved. Dr. Walter has set up his own laboratory at the University of California, San Francisco, and has recently made major strides in determining the structure and function of the SRP. His group has successfully disassembled SRP and purified its individual components to homogeneity. They found that SRP consists of:

1. One molecule of 7SL RNA, 300 nucleotides long. Approximately 100 nucleotides at the 5' end and 45 nucleotides at the 3' are homologous to the human Alu right monomer sequence (Alu is a family of DNA sequences repeated approximately 10,000 times in the human chromosome). The central S segment of 155 nucleotides shows no homology to the Alu family.
2. Four different proteins:
   19 kd monomer
   54 kd monomer
   heterodimer of 9 kd and 14 kd polypeptides
   heterodimer of 68 kd and 72 kd polypeptides

Overall, SRP has a molecular weight of 325,000 kd. Electron microscopy using negative stain, platinum shadowing, and darkfield showes SRP as a rod-like structure 24 nm in length and 5 nm wide. The RNA molecule spans its full length as a backbone.

SRP has three known functions: signal recognition, elongation arrest, and translocational promotion. Using biochemical dissection and reconstitution in which a single peptide is modified, as well as chemical crosslinking, Dr. Walter was able to define these functional determinants along the molecule. The elongation arrest function of SRP is contained in a separate and separable domain that consists of the two smallest SRP polypeptides (9 kd and 14 kd) and those stretches of 7SL RNA that are homologous to the repetitive Alu DNA sequence family. The four remaining SRP polypeptides (19 kd, 54 kd, 68 kd, 72 kd) together with the unique middle stretch of 7SL RNA form an SRP subparticle that is sufficient for signal recognition and proper targeting of nascent proteins to the ER membrane. Photocrosslink experiments revealed that the 54 kd peptide contains the SRP signal peptide binding site.

Dr. Walter has accomplished a major technical tour de force in dissecting the components of the SRP ribonucleoprotein complex and in determining their functional relationships. As a further refinement of this work, he has begun cloning and sequencing the genes for each SRP component. He has developed a mathematical simulation of the kinetic pausing and effects of SRP on translation and translocation. And he has successfully purified and begun the characterization of the ER membrane SRP receptor. In recognition of the high quality and significance of his contributions, Dr. Walter has been awarded the prestigious Eli Lilly & Company Research Award by the American Society for Microbiology for 1988.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

Memorandum

Date August 23, 1988

From Program Administrator, CMBD Program, NIGMS

Subject RESEARCH HIGHLIGHT: "The Establishment of Hepatocyte Cell Surface Polarity"
GM 29185, (Hubbard, Ann L.) Johns Hopkins University, Baltimore, Maryland

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

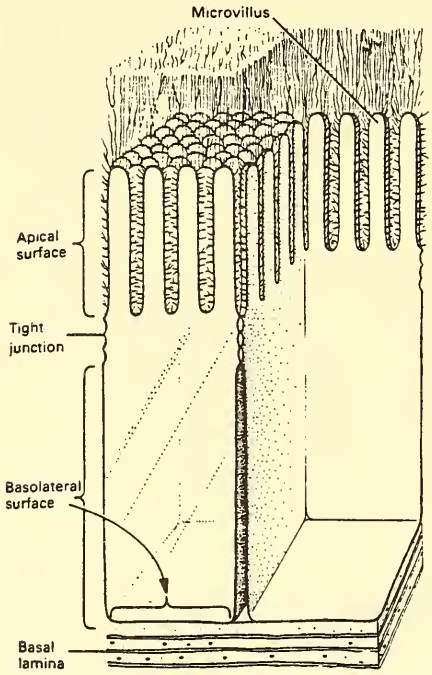


Figure 1. The Structure of Liver Parenchymal Epithelial Cells

The mechanisms by which integral membrane proteins are sorted to distinct organellar locations have received considerable attention in comtemporary cell biology. NIGMS grantee Dr. Ann Hubbard (GM 29185), Johns Hopkins University, has focused on one aspect of this central problem: how does a polarized epithelial cell establish and maintain specialized morphological and functional domains with distinct protein compositions? Her model system has been the principal epithelial cell of adult liver, the hepatocyte. As Figure 1 illustrates, hepatocytes form a cellular barrier that separates the internal (blood) and external (bile) environments. The apical or bile canalicular domain is specialized for bile secretion and is characterized by numerous microbilli. The basolateral domain includes both the lateral membrane, which is involved in cell-cell adhesion and communication and thus is marked by junctional complexes, and the basal or sinusoidal membrane, which is specialized for the exchange of metabolites with the blood. The hepatocyte must express and maintain this polarity in the face of on-going biogenesis, turnover, endocytosis, and re-cycling of each domain's membrane constituents.

However, neither the extent to which each of these processes occurs in differentiated hepatocyte membrane domains, nor their pathways, nor themechanisms that maintain the specificity of the various membranous compartments in the pathways, are yet known at the molecular level.

The biosynthetic pathway traversed by individual proteins destined for different hepatocyte cell surface domains is shared with other molecules. This requires that selective segregation and concentration occur to insure correct delivery of the various products. The pathway involves multiple, successive fusions between discrete membrane compartments, such as the endoplasmic reticulum, Golgi apparatus, and plasma membrane, thus imposing another level of sorting to insure that membrane specificities are maintained. To dissect these complex pathways, Dr. Hubbard began by preparing poly- and monoclonal antibodies against specific adult rat hepatocyte plasma membrane proteins. She developed ingenious cell fractionation procedures, using a combination of free flow electrophoresis and sucrose density gradient centrifugation, in order to purify individual proteins and unambiguously determine their immunolocalization patterns. In several cases she was able to determine the enzymatic activity or function of these domain markers:

Table 1. Hepatocyte Cell Surface Domain Marker Proteins

Protein	Domain	Function
HA16	uniformly distributed	integral membrane protein
HA321	lateral	desmosomal sialoglycoprotein
HA4	apical	?
HA201	apical	leucine aminopeptidase N
HA301	apical	dipeptidyl peptidase IV
C9	basolateral	?
ASGP-R	basolateral	asialoglycoprotein receptor
EGF-R	basolateral	epidermal growth factor receptor

With these marker reagents in hand, Dr. Hubbard used *in vivo* pulse-chase metabolic radiolabelling to compare the post-translational pathways taken by these domain-specific integral membrane proteins. She found that both apical and basolateral proteins reach the plasma membrane with similar kinetics, and achieve maximal specific activity after only 45 minutes of chase. Next, Dr. Hubbard examined the domain localization of these newly synthesized plasma membrane proteins. Surprisingly, both classes of nascent proteins were found first in the basolateral domain fraction. Newly synthesized apical proteins only began to leave this compartment after 45 minutes of chase, and reached the apical domain at different rates. HA201 and HA301 had an approximate half-time of arrival of 90-120 minutes, while HA4 moved substantially slower ($t1/2 > 150$ minutes).

These results suggest a mechanism for hepatocyte plasma membrane biogenesis in vivo in which every integral membrane protein is shipped first to the basolateral domain. Apical proteins are then subsequently re-routed to their mature, functional location. All available evidence supports the notion that tight junctions restrict lateral protein diffusion between basolateral and apical domains. Therefore, it is likely that the reassortment of nascent apical proteins results from selective vesicle-mediated endocytosis and transport. The hepatocyte has no constitutive apical secretory pathway. However, other polarized epithelial tissues (MDCK cells, intestinal epithelia) appear to use the trans-Golgi as a sorting site and acidification as a mechanism for segregating molecules into two different direct secretory pathways to both surfaces. There is precedence, however, for selective basolateral-to-apical movement in the hepatocyte. Transcyotic receptors for ligands such as polymeric IgA and hemoglobin-haptoglobin complexes have a functional life cycle which involves binding a circulating ligand at the basolateral surface, endocytotic vesicle transportation to the apical pole, fusion with the apical plasma membrane and proteolytic release into the bile. It seems likely that the hepatocyte would use this basolateral-to-apical transcytotic route for the delivery of all plasma membranes to their apical domain.

Dr. Hubbard now plans to use cell fractionation procedures to isolate and characterize the putative transcytotic vesicles. She also will examine metabolic perturbation (such as fasting or treatment with drugs) which might affect this transportation. This work should help clarify whether polymeric IgA and nascent apical proteins travel together from blood to bile and reveal the molecular details of this traffic.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

gb

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 6, 1988

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlights: "Idiotypes Bring Out the Receptor"
GM27155 (Blobel, G.), Rockefeller University

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Dr. Gunter Blobel has long been the key player in the field of protein trafficking in the eukaryotic cell. He and his group have provided the theoretical framework which has guided work in this area. And he has consistently developed biochemical and molecular genetic approaches to validate the detailed predictions of his hypotheses. In its simplest form, his working model consists of two propositions: First that nascent polypeptides contain particular amino acids--variously termed signal, transit, or targeting sequences--which contain topologic information. These sequences direct the protein to its appropriate mature cell compartment. Second, during topogenesis the targeting sequences (together with possible soluble accessory factors) interact with membrane-bound receptors as the protein is "captured" to its functional, final location.

Dr. Blobel has devoted a considerable body of work to the demonstration and characterization of topologic sequences in many proteins. Using in vitro mutagenesis, recombinant DNA production of chimeric proteins, and cloning and sequencing strategies, he has identified the location and primary structure of signal sequences which direct various families of polypeptides to such locations as the endoplasmic reticulum, plasma membrane, mitochondrion, lysosome, peroxisome, nuclear envelope, or chloroplast. He has characterized cytoplasmic factors which interact with these targeting sequences and described their physiology. Comparative studies of polypeptides which are sorted to the same compartment have revealed many conserved features. For example, all secretory proteins which are sorted into the endoplasmic reticulum contain short NH_2-terminal hydrophobic signal sequences which are removed by proteolysis during capture. Proteins destined for the nuclear envelope, on the other hand, carry longer COOH-terminal targeting sequences which remain in the mature polypeptide. Interestingly, primary sequence analysis has revealed that those topogenic sequences specific for a single organelle are not homologous to each other--although they may share similar three-dimensional structure or net chemical character. This observation has now focused considerable attention on the second half of Dr. Blobel's model.

The study of putative protein localization receptors has lagged far behind the demonstration of targeting sequences. Organelles are difficult to purify and the solubilization and biochemistry of their integral membrane proteins is even more problematic. However, in the past year, Dr. Blobel has reported a new technical approach which has resulted in the first identification and characterization of a protein targeting membrane-bound receptor. His model system is the pea chloroplast. The small subunit (S) of ribulose-1, 5-bisphosphate carboxylase is synthesized on free ribosomes as a larger precursor (pS) containing an NH_2-terminal transit peptide. Dr. Blobel had developed an in vitro assay for the postranslational import of pS into isolated chloroplasts, accompanied by cleavage of the transit sequence. To identify the putative receptor, Dr. Blobel turned to the idiotype network theory of Jerne. Jerne's concept proposes that idiotypes (or antigenic determinants associated with the antibody's antigen-binding site) can act as immunogens. Antibodies which recognize idiotypes are referred to as anti-idiotype antibodies. Dr. Blobel reasoned that among antibodies generated against the transit sequence would be antibodies which recognize the transit sequence in a way that mimics the physiological receptor-ligand interaction. Consequently, the idiotypes of some of the antibodies against the ligand may have structures in common with the receptor. A second set of antibodies raised against "receptor-like" idiotypic determinants of the anti-transit peptide antibodies would contain anti-idiotypic antibodies which recognize the ligand-binding site of the receptor.

Thus, Dr. Blobel raised antibodies against a chemically-synthesized peptide representing a 30 residue COOH-terminal portion of the pea pS transit peptide. Then, he raised a second set of antibodies against the anti-transit sequence antibodies. These presumptive anti-idiotypic antibodies caused aggregation and precipitation of isolated chloroplasts, but did not react with pS. When Fab fragments of the anti-idiotypic antibody were included in the in vitro assay, they completely inhibited pS import. Immunofluorescence and immunogold studies revealed that the anti-idiotypic antiserum recognized a determinant on the periphery of the chloroplast, in patches which coincide with contact zones where outer and inner envelope membrane are apposed to each other. The anti-idiotypic antibodies specifically recognized a 30 K integral membrane protein, judged by Western blot analysis.

Dr. Blobel's strategy of anti-idiotypic antibodies can now be generalized to other membrane receptors. This advance should lead to a very rapid increase in our knowledge of the molecular details of the protein targeting process.

Lore Anne McNicol

Lore' Anne McNicol, Ph.D.

Memorandum

Date May 26, 1989

From Program Administrator, CMBD, NIGMS

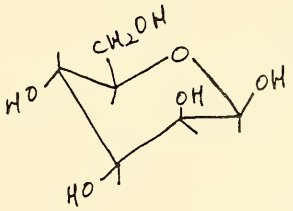
Subject Research Highlight: "Recycling of Cell Surface Mannose-6-Phosphate Receptor Molecules" 1 R01 GM 38183 (Snider, M.D.) Case Western Reserve University

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Proteins have been the charismatic biopolymer of the 80's. They perform much of the cell's biochemical work, and the genetic wizards can design and alter their amino acids with breath-taking ease. Therefore, a lot of time and venture capital has been focussed on turning proteins into pharmaceuticals. Nevertheless, there are other biochemical fish to fry. Carbohydrates also control many of the cell's functions and form a significant portion of the structure of mature proteins. Some biotechnologists are beginning to switch allegiance to begin tinkering with sugars.

Polysaccharides are intrinsically more complex than proteins: Three amino acids can be arranged in a peptide chain in six different ways, while three hexoses can be arranged into 980 different oligosaccharides. Such sugars appear on glycoproteins for more than decoration. In recent years it has become clear that sugars can direct proteins to particular tissues or cells in the body, can affect protein stability and solubility, can enhance or dampen the catalytic activity of proteins, and protect proteins from destruction. NIGMS grantee Dr. Martin D. Snider, Case Western Reserve University, has been actively studying the function of N-terminal mannose-6-phosphate (M6P) residues (see Figure 1) on glycoprotein trafficking in the cell. This sugar is known to be used as a marker for lysosome biogenesis:

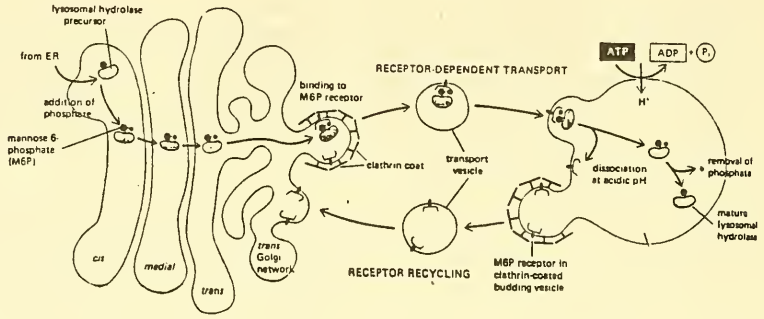
Figure 1 - Mannose-6-phosphate



These organelles contain about 40 distinct acid hydrolases (proteases, nucleases, lipases, glycosidases, phospholipases, phosphatases, and sulfatases) which are used for the controlled intracellular digestion of

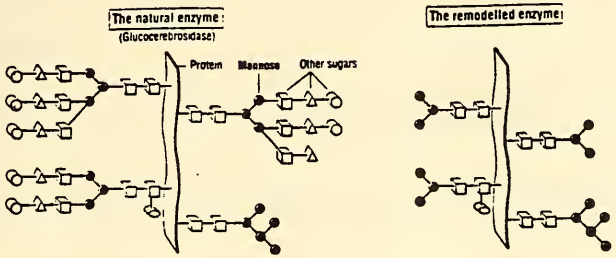
macromolecules. The precursor forms of these lysosomal hydrolases are tagged with M6P in the cis Golgi and segregated from all other types of protein in the trans Golgi network (see Figure 2). This segregation can occur because clathrin-coated vesicles budding from the trans Golgi concentrate M6P-specific receptors. These receptors are transmembrane proteins which bind the sugar-marked lysosomal hydrolases. This use of an oligosaccharide marker for "cargo recognition" by a shuttling vesicular transport system is the best model available for understanding the many sorting events that occur in a enkaryotic cell.

Figure 2 - Golgi apparatus



Dr. Snider discovered that the M6P receptor can also be found in the plasma membrane, where it functions in receptor-mediated endocytosis of lysosomal enzymes which have been released into the circulation at other sites. This scavenger pathway recaptures any hydrolases which have escaped the normal packaging process. Using cell fractionation and immunoprecipitation techniques, Dr. Snider found that 10-20% of the total pool of M6P-receptor is on the cell surface. The surface receptor molecules turn over rapidly and mix with the Golgi M6P-receptor pool with a t1/2 = 1-2 hr.

Dr. Snider's observations are finding application to the treatment of Gaucher's disease, a rare and potentially fatal inherited disorder which causes the spleen to enlarge to the size of a basketball and distorts bones. Children suffering from Gaucher's disease lack the enzyme cerebrosidase, which breaks down cerebroside lipids in macrophages. A treatment trial with purified natural enzyme had not been successful because the diseased macrophages could not internalize the exogenous cerebrosidase. However, this year the biotechnology firm Genzyme has treated the enzyme so that its terminal oligosaccharide residues are mannose-6-P:



The remodeled enzyme is now a substrate for Dr. Snider's cell surface M6P receptor. Eight treated patients have shown both marked reduction in their splenomegaly and restored bone growth.

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 23, 1989

From Program Administrator, CMBD Program, NIGMS

Subject Research Highlight: "Molecular Mechanisms of Intracellular Translocation and Metabolism of Lipids" ROI GM37434 (Pagano, R.E.), Carnegie Institution of Washington

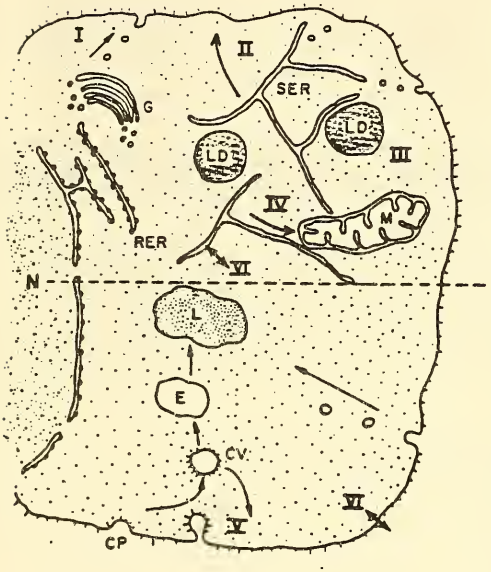
To
Director, NIGMS
Through: Director, CMBD Program, NIGMS

Plasma membranes are composed primarily of proteins and phospholipids. While our knowledge of the biosynthesis and insertion of membrane proteins has been greatly expanded in recent years, little is known about the processing of membrane phospholipids. In bacteria, phospholipid molecules are synthesized at the inner leaflet of the plasma membrane. After a short lag, they equilibrate with the outer leaflet of the lipid bilayer. In eukaryotic cells, the situation is much more complex. Most phospholipid synthesis occurs at the endoplasmic reticulum, and then individual lipid molecules are somehow distributed into all intracellular membranes. Different lipid species are present in different amounts in the various intracellular organelles, and the species of lipid are not uniformly distributed on the two halves of the membrane bilayer. Since different lipids may have dramatically different fatty acid compositions, this lipid asymmetry produces membranes in which the two leaflets have markedly different fluidity properties, and presumably, different functions.

Lipids pose unique problems to the cell biologist which makes the study of their intracellular distribution and transport particularly difficult. For example, labelled lipids rapidly exchange among intracellular membranes during standard cell fractionation procedures. Therefore, localization studies must be performed in situ. Also, in contrast to studies of protein translocation which utilize labelled and specific antibodies, hormones, or toxins, no specific probes for the major classes of lipids exist. NIGMS grantee Dr. Richard E. Pagano, Carnegie Institution of Washington, has made a major methodological advance in the study of lipid translocation by developing a low-light-level digital imaging fluorescence microscopy system. Dr. Pagano's group has synthesized a variety fluorescent-labelled lipid species, using C_6 - nitrobenzodiazole derivatives. Their general strategy has been to prepare carrier lipid vesicles containing the appropriate fluorescent analog. These vesicles are incubated with tissue culture cells for 90 min at 2 C (a condition which prevents adsorption, endocytosis, or fusion and permits a direct lipid transfer process) and then washed and examined under physiological conditions.

As diagrammed in Figure 1, workers in the field had postulated six possible molecular mechanisms that could direct lipid traffic:

Figure 1. Possible Molecular Mechanisms of Lipid Trafficking



CP	coated pit
CV	coated vesicle
E	endosome
G	Golgi apparatus
L	lysosome
LD	lipid droplet
M	mitochondrion
N	nucleus
RER	rough endoplasmic reticulum
SER	smooth endoplasmic reticulum

I vesicle fusion and budding

II cytosolic transport of lipid monomers

III localized synthesis by specific organelle

IV lateral diffusion between organelles connected by membrane bridges

V cyclic translocation

VI protein-mediated transmembrane flip-flop

Dr. Pagano's fluorescence techniques have provided direct evidence for the existence of two of these pathways. The simple lipids, phosphatidylserine and phosphatidylethanolamine, use mechanism VI. At 2 C, both fluorescently-labelled lipids rapidly entered the plasma membrane and translocated into the cytosol. This process required ATP and was inhibited by N-ethylmaleimide, suggesting a protein-mediated, energy-dependent process. In contrast, phosphatidyl choline and a variety of ceramides were shown to utilize mechanism I. At 2 C these species remained in the plasma membrane. Upon warming to 37 C, both were internalized to the Golgi apparatus by vesicular endocytosis. In an intriguing observation, Dr. Pagano noted that a simultaneously bound fluorescent-labelled protein lectin segregated from these lipids during the internalization process. Therefore, the trafficking apparatus appears to discriminate between proteins and lipids.

Thus, Dr. Pagano's findings on fluorescent lipid trafficking are striking and much different from what one might have anticipated from knowing the protein internalization pathways. He plans to continue to explore the molecular basis for these mechanisms.

Lore' Anne McNicol, Ph.D.

LAMcNicol/gb

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 13, 1988

From Program Administrator, CMBD, NIGMS

Subject Research Highlight: "Role of Cell Adhesion Molecules in the Formation and Maintenance of Epithelial Cell Junctional Complexes"
ROI GM37432 (Gumbiner, B.M.), UCSF

To Director, NIGMS
Through: Director, CMBD, NIGMS

In vertebrate animals, epithelial tissues form selective, polarized barriers between the lumenal and serosal organ spaces. Epithelia generate and maintain large ionic and molecular compositional differences between these compartments in order to function in vectorial transport processes. The intracellular junction (diagrammed in Figure 1) is one of the major hallmarks of absorptive and secretory epithelia. It contributes to the permeability barrier by controlling diffusion through the paracellular space, acting as a "gate" with highly variable permeability properties. This junction is also crucial to the organization of the transcellular transport route, functioning as a "fence" to form the boundary which distributes ion channels, pumps, and enzymes into apical and basolateral domains.

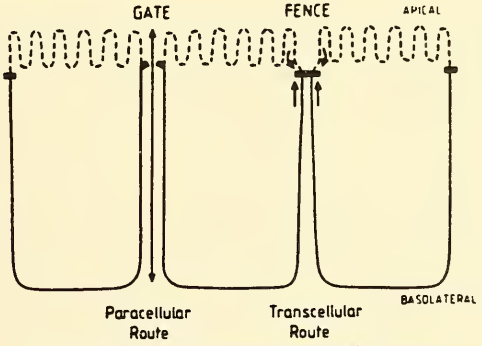


FIG. 1. Gate and fence functions of epithelial tight junctions.

NIGMS grantee Dr. Barry M. Gumbiner, University of California, San Francisco, has been studying the molecular organization of epithelial cell junctional complexes and the mechanisms by which they are assembled. His model system is

Madin-Darby canine kidney (MDCK) cells, a polarized epithelial cell line isolated from the kidney of a cocker spaniel. When cultured on permeable supports, these cells will grow to form a polarized monolayer which is impermeable to ions, has measurable transepithelial electrical resistance, and has several defined apical and basolaterial surface markers. The well-studied ultrastructure of these monolayers is outlined in Figure 2:

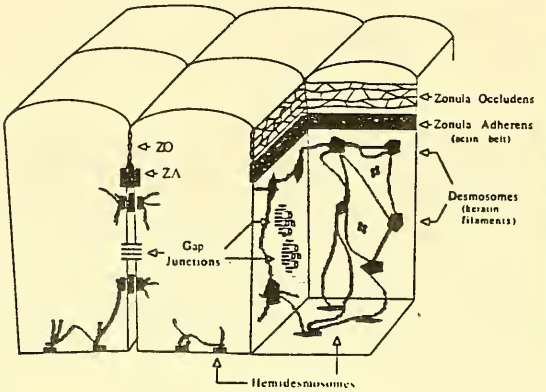


FIG. 2. Epithelial junctional complex. Zonula occludens (ZO) and zonula adherens (ZA) both form continuous zonular structures circumscribing cell apex. ZO is at boundary between apical and basolateral plasma membrane domains. ZA lies just below the ZO and is major site of attachment of actin filament bundles to lateral plasma membrane. Desmosomes are macular (spot-like) junctions found all over basolateral membrane, sometimes at greater frequency near zonular junctions and sometimes at basal surface as hemidesmosomes. They are sites of attachment of cytokeratin intermediate filaments and presumably play a role in structure and integrity of the epithelium as a whole. Gap junctions are not traditional members of epithelial junctional complex (14) but are quite common in many epithelia. They allow metabolic communication between cells by providing channels for diffusion of ions and small molecules between neighboring cells.

The junctional complex consists of three morphologically distinct regions: the zonula occludens (ZO) or tight junction, the zonula adherens (ZA) or intermediate junction, and the desmosome. The ZA is characteristically attached to cytoskeletal actin microfilaments, while the desmosomes are associated with keratin intermediate filaments. Dr. Gumbiner developed methods to isolate junctional complexes and prepare monoclonal antibodies (mabs) against junctional molecules. He also set up a sensitive electrophsiological assay to test the presence of a functional occluding barrier in MDCK cell cultures. This has enabled him to detect those molecules which participate in establishing the epithelial junction. He found that certain of his mabs would specifically inhibit the formation of active junctions and he was able to identify their target molecule as uvomorulin. Uvomorulin is a glycoprotein which, as its name suggests, functions during embryonic development. It mediates the Ca++-dependent cell adherence phenomena which lead to compaction of the morula stage embryo.

In addition, uvomorulin is a member of one of three related families of Ca++-dependent cell adhesion molecules, or "cadherins":

Cadherin Families

all cadherins are transmembrane glycoproteins which exhibit Ca++-dependent changes in molecular conformation during cell-cell adhesion, three different families identified to date are:

E	N	P
E-cadherin (epithelial)	N-cadherin (neuronal)	P-cadherin (placenta)
L-CAM (liver)	A-CAM (cardiac discs)	
cell CAM	Ncal CAM (retina)	
uvomorulin		

The E and N families have close sequence homologies to one another, with identical N-terminal amino acid sequences. The E and P families are 58% homologous.

Dr. Gumbiner next examined whether the inhibitory action of anti-uvomorulin mabs was specific for any particular subset of junctional elements (ZO, ZA, desmosomes). He found no specificity; inhibition of uvomorulin function interfered with all three forms of cell contact. In contrast, the transepithelial electrical resistance of confluent, well-established MDCK monolayers was not affected by anti-uvomorulin antibodies. Therefore, uvomorulin seems to mediate an early adhesion event that is an hierarchical prerequisite for all further assembly. Once the junctional complex is firmly in place, uvomorulin has a much less important role in maintaining intracellular adhesion.

Although ultrastructure suggests that the tight junction is a static network of strands, it actually is a dynamic structure which must assemble and disassemble during developmental and physiological processes. Dr. Gumbiner has proposed the model outlined in Figure 3 for the role of cadherins in modulating junctional structure. He postulates that uvomorulin is randomly distributed on the surface of nonpolarized cells. As the cells attach to a supporting substrate, uvomorulin molecules re-distribute and accumulate in the region of intercellular contact and mediate cell-cell adhesion. This process mediates assembly of the ZA, which in turn, provides the contact for ZO assembly. Remodeling of the junction occurs as Ca++ removal alters uvomorulin conformation, causing dissociation of the ZA. Dr. Gumbiner will test this model more directly by identifying and characterizing other protein components

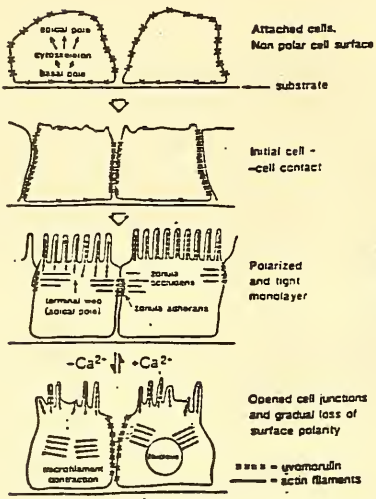


Fig. 3. Model for assembly and maintenance of epithelial tight junctions emphasizing roles of uvomorulin and the zonula adherens (ZA). Formation of ZA is initiated between nonpolar cells by uvomorulin-dependent cell-to-cell adhesion. Uvomorulin first accumulates in region of intercellular contact. Under influence of the actin-cytoskeleton, ZA assembles and becomes localized to apical region of cell. ZA provides necessary adhesive contact for assembly of the zonula occludens from tight junction strands, which otherwise would form unorganized all over lateral plasma membrane. Opening of tight junctions by removal of Ca^{2+} results from Ca^{2+} -dependent loss of uvomorulin mediated adhesion, causing dissociation of ZA.

of the tight junction and determining their molecular interactions. This work will continue to provide deeper insights into such physiological phenomena as wound healing, neutrophil transepithelial migration during inflammation, spermatogenesis, and hormonal regulation of the paracellular pathway of ion transport.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 24, 1988

From Program Administrator, CMBD, NIGMS

Subject Research Highlight: "The Diacylglycerol Phosphate Paradox: Oncogene Action or Membrane Biosynthesis." R01 GM22942 (Pagano, R.E.), Carnegie Institution of Washington

To Director, NIGMS
Through: Director, CMBD, NIGMS

Lipids have long fascinated membrane biologists and physical chemists because encoded in these simple molecules is the ability to spontaneously form macroscopic, two-dimensional membrane systems whose very shape and bulk properties depend exquisitely on the chemical makeup of the individual components. Recently, cell biologists and biochemists have become more interested in lipids due to the discovery that several classes of lipids have profound effects on cell function (phosphatidyl inositides, diacyl glycerols, eicosanoids, and ether lipids) and to the elegant description of the low density lipoprotein transport system and its regulation of cholesterol metabolism. Thus, after many years of basic studies on lipid enzymology and physical chemistry, there is a sense of renewed interest and excitement in the field of lipids.

NIGMS grantee Dr. Richard E. Pagano has refocussed our attention to the need for further fundamental studies of lipid intermediary metabolism in order to reconcile the newly appreciated regulatory functions of lipid species with their known positions in membrane biosynthetic pathways. This need is particularly acute for our understanding of the cell biology of sn-1, 2-diacylglycerol (DAG). As illustrated in Figure 1, many neurotransmitters, hormones, and growth factors appear to utilize DAG as a second messenger.

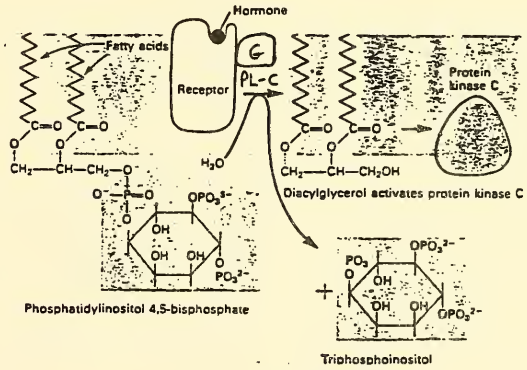


Figure 1. Second messenger function of diacylglycerol.

Receptor binding initiates transmembrane signalling (probably transduced through G-protein activation of phospholipase C) which cleaves phosphatidylinositols to DAG. DAG is freely soluble in the lipid bilayer and can activate protein kinase C, an enzyme implicated in the regulation of cell activation, differentiation, proliferation, oncogene action, and tumor promotion in a variety of cell types. Despite the limited aqueous solubility of DAG's carrying physiologically long chained fatty acids, there is considerable direct evidence for this second messenger function. Since DAG also plays an "ordinary" housekeeping role as a precursor of glycerolipids, there is potential for the existence of several independently-regulated intracellular DAG pools. Defects in DAG metabolism might alter these pools to cause cell transformation, and enzymes of DAG metabolism might be oncogene products. Obviously, the molecular mechanisms regulating DAG metabolic pathways need to be understood.

Dr. Pagano has called attention to the fact that significant amounts of DAG are produced during sphingomyelin (SM) biosynthesis. Classic work by NIGMS grantee Dr. Eugene P. Kennedy, Harvard University, had determined that this sphingolipid is produced enzymatically by transfer of a phosphorylcholine group to a ceramide (Cer), as phosphatidyl choline (PC) is converted to DAG (Figure 2).

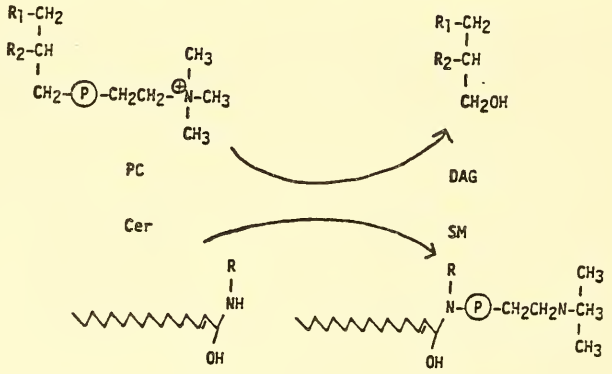


Figure 2. Biosynthesis of sphingomyelin.

Using fluorescent analogs of these reactants, Dr. Pagano was able to use low-light-level digital fluorescent microscopy to determine the site of SM synthesis, and its subsequent fate. Upon an early pulse he found SM to be intensely localized at the Golgi apparatus, constituting 15% of its lipidic mass. After synthesis, the SM was rapidly transported to the plasma membrane. It is obvious that a molecule of DAG is generated for each SM synthesized. Dr. Pagano then worked backwards, using similar methodology to observe that the PC and Cer precursors are themselves produced in the endoplasmic reticulum (ER). In fact, PC is the most abundant lipid constituent of the ER, comprising nearly 70% of the total lipid in that organelle. Citing recent work from Dr. James E. Rothman's group at Princeton which defined a very extensive vesicular transport from the ER to the Golgi apparatus, Dr. Pagano has pointed out the need for "recycling" of lipid monomers back to the ER in order to maintain its structure. He has proposed the following model, using DAG as a possible candidate for such a lipid recycling process:

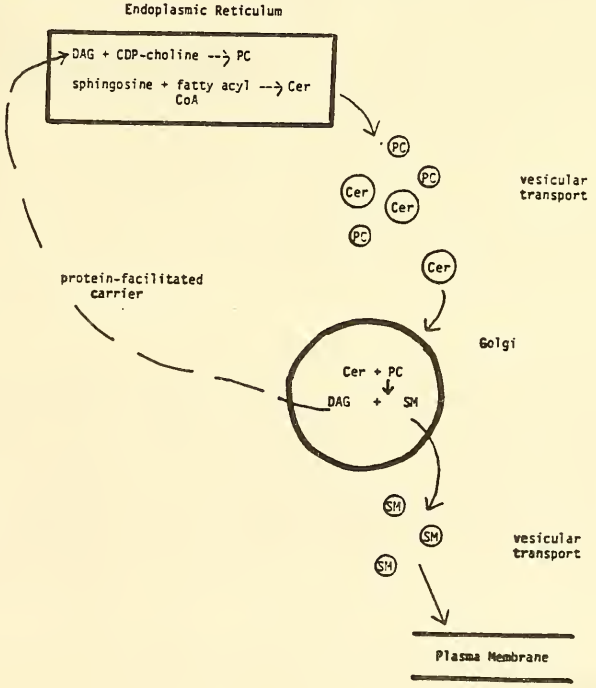


Figure 3.

His model makes several specific predictions:

1. There exists a facilitated carrier which moves DAG from the Golgi to the ER.
2. DAG would accumulate in enriched Golgi fractions when exogenous ceramide was supplied but would subsequently disappear with further addition of the carrier and a membrane acceptor.
3. Disruption of ER to Golgi vesicular traffic would decrease cellular SM levels.

In preliminary experiments, Dr. Pagano has observed facilitated transport of exogenously supplied DAG fluorescent analogs. He is presently attempting to isolate and characterize the transporter, and test other aspects of his recycling model. Thus, Dr. Pagano has revitalized this traditional lipid biochemistry field, and brought it to a position to make exciting contributions to studies of signalling and protein trafficking.

Lore' Anne McNicol, Ph.D.

LAMcNicol/gb

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 11, 1989

From Program Administrator, CMBD, NIGMS

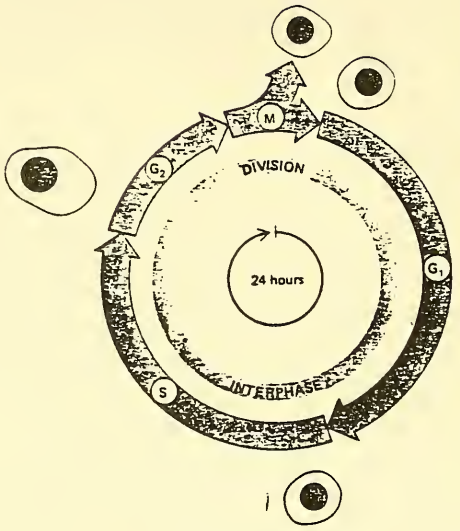
Subject RESEARCH HIGHLIGHT: "Overview of Cell Cycle Control" P01 GM31286-07 and R01 GM39023-02 (Kirschner, Marc W.), University of California, San Francisco; R01 GM39620-02 (Beach, David H.), Cold Spring Harbor Laboratory, New York; R01 GM19363-25 (Gerhart, John), University of California, Berkeley; R01 GM26743-10 (Maller, James L.), University of Colarado Health Sciences Center, Denver, Colorado; R01 GM40658-02 (Lohka, Manfred J.), University of Colorado Health Sciences Center, Denver, Colorado.

To Director NIGMS
Through: Director, CMBD Program

Cell division is one of the critical events in the life cycle of a cell. In the past year, major strides have been made towards understanding the biochemical events which drive the cell cycle through division. This understanding has come from the convergence of studies being conducted in multiple laboratories around the country and abroad, using species ranging from yeast to man. Many of the key players in this emerging story have been supported by NIGMS in the CMBD program. Following is an overview of how these disparate lines of research, from the study of cell division genes in yeast, to purification of maturation promotion factor in frog eggs, to characterization of cell cycle control in mammalian cells, have come together to reveal a common biochemical machinery which controls cell division. This overview will be followed by three individual research highlights detailing the purification of maturation promotion factor (MPF), and the biochemical and functional characterization of its major component proteins, a p34 kinase and cyclin, in cell cycle regulation.

The story begins with MPF, which was originally described in 1971 in immature frog eggs. It was found to promote meiosis and maturation of the immature eggs, and also to induce mitosis. In the early 1980's Drs. Gerhart and Kirschner observed that MPF activity fluctuated during the cell cycle, peaking as cells entered mitosis and dropping abruptly as cell division was completed (see figure 1 for a diagram of the cell cycle.) Unfortunately, purification of MPF, which was known to be a protein, proved particularly difficult, and was not accomplished until 1988 when James Maller and Fred Lohka devised an in vitro system to assay MPF activity. They then discovered that MPF was in fact a complex of two proteins, one of molecular weight 34,000 and the other of molecular weight 45,000.

Figure 1.



The four successive phases of a typical eucaryotic cell cycle. After the *M phase*, which consists of nuclear division (*mitosis*) and cytoplasmic division (*cytokinesis*), the daughter cells begin interphase of a new cycle. Interphase starts with the G_1 *phase*, in which the biosynthetic activities of the cells, which proceed very slowly during mitosis, resume at a high rate. The *S phase* begins when DNA synthesis starts, and ends when the DNA content of the nucleus has doubled and the chromosomes have replicated (each chromosome now consists of two identical "sister chromatids"). The cell then enters the G_2 *phase*, which continues until mitosis starts, initiating the M phase. During the M phase, the replicated chromosomes condense and are easily seen in the light microscope. The nuclear envelope breaks down (except in some unicellular eucaryotes such as yeasts, where it remains intact), the sister chromatids separate, two new nuclei form, and the cytoplasm divides to generate two daughter cells, each with a single nucleus. Cytokinesis terminates the M phase and marks the beginning of the interphase of the next cell cycle. A typical 24-hour cycle is illustrated here, although cell-cycle times in eucaryotic cells vary widely, from less than 8 hours to more than a year in adult animals, with most of the variability being in the length of the G_1 phase.

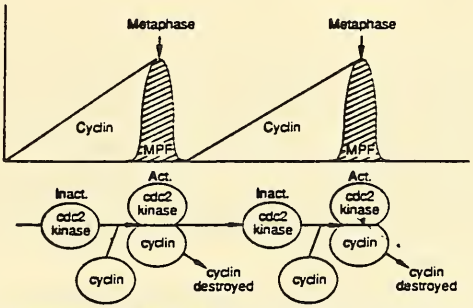
il Growth and Division

In the meantime studies being conducted by Paul Nurse in England and David Beach at Cold Spring Harbor, had identified a gene in fission yeast which was required for division, termed cdc-2. This gene encoded a protein kinase with a molecular weight of 34,000. David Beach subsequently identified a kinase in mammalian cells which was highly homologous to the yeast protein and which in fact could replace the cdc-2 gene product in yeast mutants lacking cdc-2. This p34 kinase turned out to be identical to the 34,000 MW component of MPF. About the same time that MPF activity had been described in the frog, several investigators demonstrated a protein in sea urchin and clam embryos which had a regulatory role in the cell cycle. It was called cyclin because it rose and fell with the cell cycle and in fact could trigger meiosis in maturation-arrested frog oocytes. A 45,000 MW cyclin homolog has now been shown to be the second component of frog MPF, and a 13,000 MW homolog also is found complexed to the cdc-2 protein kinase in yeast. Thus cell division in yeast, sea urchins, clams, frogs, and man is under the control of two proteins, a kinase

and a cyclin, which oscillate in activity and content, respectively, with the phases of the cell cycle.

The most recent findings from the laboratories of Marc Kirschner and Andrew Murray, and David Beach, are beginning to reveal how these two components of MPF interact and regulate progression through the cell cycle (see figure 2). Briefly, cyclin levels rise to a peak at metaphase. When a critical cyclin concentration is reached, the kinase component of MPF is activated, at which point cell division begins; the cyclin component is then rapidly degraded, and the kinase becomes inactivated.

Figure 2.



Controlling the cell cycle. *Cyclin concentrations increase until the cdc2 kinase component of maturation promoting factor (MPF) is turned on at metaphase, the midpoint of mitosis. The cyclin is then degraded and MPF activity drops as the cell divides. Then another cycle starts. [Adapted from G. Draetta et al., Cell 56, 829 (1989).]*

While this unified picture of cell cycle regulation has emerged, many questions remain to be answered. How does cyclin act (directly or indirectly) to activate the p34 kinase; what is the substrate for the kinase; how is cyclin degraded? The answers to these questions are already on the way and should provide the fine details to the larger picture of a universal cell cycle control mechanism which appears to operate in species ranging from the simple yeast, to man.

Marion Zatz, Ph.D.

Reference

Marx, J.L. The Cell Cycle Coming Under Control. Science, 245, 252-255, 1989.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 18, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "Cell Cycle Control I - Maturation Promotion Factor Comes of Age" R01 GM 26743-10 (Maller, James), and R01 GM 40658-02 (Lohka, Manfred), University of Colorado, Health Sciences Center, Denver Colorado

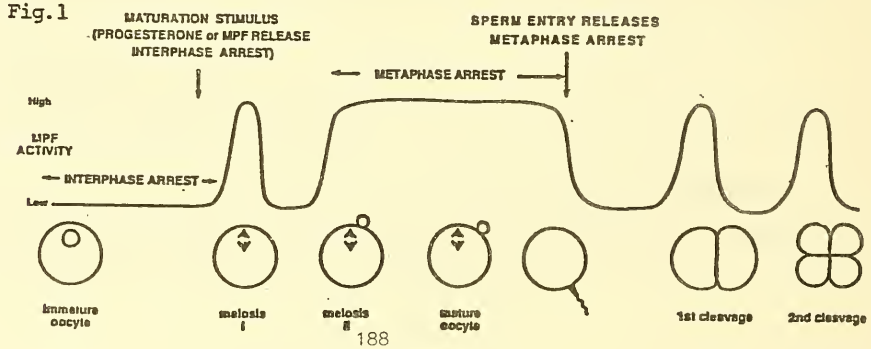
To

Director, NIGMS
Through: Director, CMBD

The frog oocyte and egg provide an important model system for the study of the regulation of the cell cycle. The cell cycle is defined as the set of events responsible for the duplication of the cell, the basic unit of life. Replication of the genetic material occurs in the S (synthetic) phase and segregation of the replicated DNA, destined for each daughter cell, occurs in the M (mitotic) phase. Usually a G_1 phase separates M from S phase, and a G_2 phase separates S from M phase. In the egg, there are key points at which the cycle is physiologically arrested; the studies of the factors which overcome these arrest points are revealing the underlying biochemical mechanisms which control the cell cycle in frog eggs, as well as somatic cells of species ranging from yeast to man.

The major cell cycle events involved in frog oocyte maturation, including natural arrest points, are depicted in Figure 1. The immature oocyte is arrested at the G_2 phase of the cell cycle, prior to the first meiotic division. Secretion of progesterone induces progression through the first meiotic division to a second arrest at the G_2 phase preceding metaphase of the second meiotic division. Fertilization overcomes the second meiotic arrest and triggers a series of mitotic divisions of the egg giving rise to the early embryo. Thus oocytes undergo a simple cell cycle comprising a $G_2 \rightarrow M$ transition at maturation, and an $M \rightarrow G_1/S$ transition at fertilization. The first meiotic division does not require protein synthesis, however the second meiotic division and all subsequent mitotic cycles do require newly synthesized proteins.

Fig.1



A major breakthrough in dissecting the biochemical control of the cell cycle came in 1971 when Masui and Markert discovered a cytoplasmic activity in the frog egg which could trigger meiotic maturation when microinjected into immature oocytes arrested in G_2. This activity, which prepares the egg for fertilization, was named maturation promotion factor (MPF), and also has been termed M-phase promotion factor because of its activity both in meiosis of germ cells and mitosis of somatic cells. MPF activity, as defined above, has been detected in mitotic cells from many species, ranging from yeast to man. Its activity oscillates with the cell cycle, becoming maximally active in late G_2, and disappearing in early G_1; thus it may be a fundamental regulator of the cell cycle.

Once this activity was discovered, several labs attempted to purify and characterize MPF; this proved to be a particularly difficult task, due to the instability of the MPF activity and the absence of a good assay system. However, in the last year James Maller and Manfred Lohka finally have succeeded in purifying MPF and defining its key components. One of the components has been determined to be a homolog of the cdc-2 gene product which regulates cell cycle in fission yeast. This accomplishment has led to an explosion of progress in the understanding of how the cell cycle is regulated in all eukaryotic cells.

Studies on MPF in the early 1980's, mainly in the laboratories of Dr. John Gerhart (R01 GM19363), University of California, Berkeley, and Dr. Marc Kirschner (P01 GM31286), University of California, San Francisco, revealed that MPF activity fluctuated with the cell cycle; frog oocytes arrested prior to the first meiotic division did not require new protein synthesis for expression of its activity, implying a post-translational control of MPF activity. Another key biochemical event associated with MPF activity was phosphorylation of cellular proteins such as histone and nuclear lamins; these findings suggested that MPF might either be a kinase or perhaps activate a kinase. With the development of a crude *in vitro* assay, in which MPF preparations were added to somatic nuclei mixed with oocyte cytoplasm, Drs. Gerhart, Wu and Kirschner were able to achieve a partial purification of MPF and define some of its properties. The assay was based on the fact that MPF caused nuclear envelope breakdown, chromosome condensation and spindle formation, some of the distinctive features of entry into M phase of the cell cycle. They learned that an inactive form of MPF, termed pre-MPF was stored in the frog oocyte, could be "auto-activated" by a small amount of partially purified MPF, and that the activity could be stabilized by thio-ATP and phosphatase inhibitors, implicating a phosphorylation reaction in MPF activity. However, due to the likely presence of MPF inactivating factors in the oocyte cytoplasm, further progress on MPF purification was not achieved until Drs. Maller and Lohka, using an improved in vitro system of demembranated interphase sperm nuclei in cell-free extracts, were able to monitor MPF purification. In the system which they developed, nuclei could go through one complete mitotic cycle. A 3500 fold purification of MPF was achieved with six chromatographic steps, and the purified fractions consistently were able to induce not only metaphase nuclei in vitro, but nuclear membrane breakdown and meiotic progression when microinjected into oocytes, *in vivo*.

This activity was associated with two proteins of 32kDa and 45kDa molecular weight, respectively. These fractions also contained a protein kinase activity which resulted in the phosphorylation of the endogenous 45kDa band, as well as histone H1. In a critical series of experiments, Drs. Maller and Lohka went on to demonstrate that antibodies to the cdc-2 p34 protein kinase could immunoblot and immunoprecipitate the 32kDa component of MPF, indicating homology to the 34kDa product of the yeast cdc-2 gene which controls entry of yeast cells into mitosis. This yeast mitotic regulator is a kinase and exists in a complex with a second protein, a cyclin homolog, of 13kDa MW. Dr. Maller speculates that the 45kDa component is the frog homolog of cyclin.

MPF initially was defined as a mysterious cytoplasmic factor which promoted maturation of frog oocytes. It now appears that MPF is a universal mitotic regulator which controls the G_2-M transition in cells of species ranging from yeast to man. It is a complex of at least two proteins, one of which is a highly conserved protein kinase of approximately 34,000 molecular weight. MPF kinase activity fluctuates with the cell cycle despite the fact that the p34 protein is present at all times. These findings raise many intriguing questions regarding cell cycle control. What are the substrates for MPF; how is its activity regulated; what is the role of the second protein in the complex? Rapid progress is being made toward providing the answers to these questions. In the next two highlights of this series some of this progress will be reviewed.

Marion Zatz

Marion Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 20, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "Cell Cycle Control II - cdc-2 Kinase, A Universal Mitotic Oscillator" R01 GM34607 and R01 GM39620 (Beach, David), Cold Spring Harbor Laboratory, New York

To

Director, NIGMS
Through: Director, CMBD Program

While the biochemists were attempting to purify frog MPF and determine how it triggers M phase, the geneticists were attempting to isolate genes which regulate the cell cycle in yeast. These two approaches have converged to identify a common, highly conserved biochemical machinery that regulates cell division in all eukaryotic cells. Dr. Beach, initially working with Paul Nurse in England, discovered a gene (called cdc-2 for cell division cycle) which controlled entry into mitosis in the fission yeast, S. pombe. Mutants lacking this gene are arrested in G_2 of the cell cycle. A similiar gene, cdc-28, was isolated from budding yeast (S. cerevisiae) by Dr. Steven Reed (R01 GM38328). These two gene products are sixty-two percent homologous for amino acid content and can replace each other in yeast mutants lacking the cdc-2/cdc-28 kinase. Subsequently, Dr. Beach found a p34 mammalian homolog of cdc-2 in HeLa cells, complexed to both a 13kDa protein and a second protein of MW 62kDa. The human homolog also can rescue temperature sensitive yeast cdc-2 mutants arrested in G_2. Most recently Dr. Beach has identified a cdc-2 protein homolog in clams. This kinase, which also is a component of frog MPF and the H_1 histone kinase activity of starfish, was found in association with the clam cyclins.

The cyclins were first described by Tim Hunt (University of Cambridge, England) and Joan Ruderman (Duke University) in marine invertebrate eggs; they are proteins which are synthesized in S phase at the beginning of the cell cycle, and selectively degraded during M phase. The cyclins were given their name because of their dramatic fluctuation with the cell cycle and it was proposed that they played an important regulatory role in cell cycle control. In support of this hypothesis, it was found that clam cyclin mRNA when injected into immature frog oocytes, directed cyclin synthesis and triggered meiosis, indicating that it was related in some way to MPF. Thus, the finding of Dr. Beach that cdc-2 kinase is associated with cyclins in clam sets the stage for understanding how MPF works, and also provides a model for the complex of cdc-2 kinase with a second protein, likely to be a cyclin, in yeast, frog, and man. Dr. Beach proposes that the kinase activity of MPF, which fluctuates with the cell cycle, is regulated by the association of cyclin with the cdc-2 kinase, and it is the active complex which triggers entry into mitosis.

Although cdc-2 was first identified as a critical regulator of M phase in yeast, many of its biochemical properties have been uncovered by Dr. Beach in studies of the human homolog, p34.

The Kinase - The p34 kinase of HeLa cells is abundant throughout the cell cycle. However, in early G_1 it is unphosphorylated, and is inactive as a protein kinase. As the cycle progresses, it becomes heavily tyrosine phosphorylated, and at least one of its tyrosine residues is selectively phosphorylated by the pp60 src oncogene, tyrosine kinase. The highly phosphorylated p34 becomes associated with a p62 protein and the p34/62 complex becomes maximally active as a protein kinase in late G_2. One of the substrates for the p34 catalytic subunit is in fact the p62 endogenous component of the complex, which also may be a human cyclin homolog. In addition to its association with the p62 protein, the kinase also is complexed with a 13kDa m.w. protein which is homologous to the yeast p13 suc-1 gene product, which also associates with the yeast cdc-2 kinase. This picture of a kinase 1) whose activity and phosphorylation state fluctuate with the cell cycle, 2) which is associated with a second polypeptide, and 3) phosphorylates, as one of its substrates, its endogenous protein partner, bears striking resemblance to the characteristics of MPF identified in frog and yeast.

The Substrate - The cdc-2 kinase phosphorylates, in addition to its protein subunit, other cellular proteins; in fact one of the characteristics of mitosis is the presence of many phosphorylated protein species. The protein kinase of frog and starfish MPF phosphorylates histone H_1, and appears to be identical to the growth associated H_1 kinases that have been known for many years to appear during mitosis in a variety of organisms. The frog kinase also has been shown to phosphorylate nuclear lamins and a phosphatase inhibitor. Histone and lamin phosphorylation may play a role in chromosome condensation and spindle formation, whereas phosphorylation of the phosphatase inhibitor may provide another level of control for phosphorylation and dephosphorylation of proteins during cell cycle progression. Very recently, James Maller has identified the pp60src proto-oncogene product as a substrate for the MPF kinase. This is an intriguing observation since it was mentioned earlier that pp60src tyrosine kinase also phosphorylates the MPF kinase. These two observations provide the first biochemical information on how the src gene might function in normal and malignant growth control.

Complex Interactions

The above information on the MPF kinase provides a view of how the cell cycle may be controlled. But it also is evident that this control will be complex, regulated by the association and disassociation of key proteins (such as kinases and cyclins) as well as by a cascade of phosphorylation/dephosphorylation events. In the following highlight some of these complexities will be addressed, and in particular how the cyclins act in conjunction with the cdc-2 kinase to regulate the cell cycle.

Marion Zatz

Marion Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 20, 1989

From Program Administrator, CMBD Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Cell Cycle Control III - The Ups and Downs of MPF Activity - The Role of Cyclins" R01 GM39023 (Kirschner, Marc and Murray, Andrew), University of California, San Francisco

To

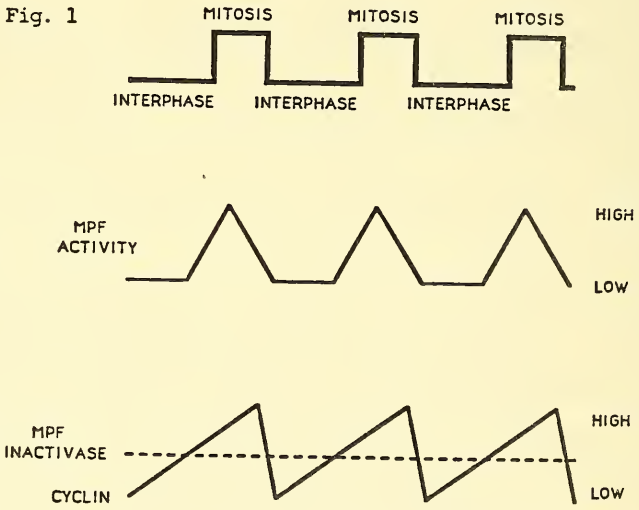
Director, NIGMS
Through: Director, CMBD Program

The previous two highlights of this series describe the purification of maturation promotion factor (MPF), and its characterization as a complex of two major proteins. One of these proteins is a 32-34kDa protein kinase which is highly conserved throughout the evolutionary scale and is homologous to the cdc-2/cdc-28 kinase of yeast. This kinase appears to be active in protein phosphorylation only in the mitotic phase of the cell cycle, although steady levels of the peptide are present throughout the cell cycle. Kinase activity seems to be dependent upon association of the kinase with a second protein ranging in M.W. from 13kDa (in yeast) to 45-65kDa in frog and man. The yeast and human cdc-2 kinase may actually be part of a tri-molecular complex with both a p13 and a p62 subunit.

The appearance of the active complex correlates both with progression of the cell cycle from $G_2 \rightarrow M$, and with phosphorylation of cellular proteins, such as nuclear histones and lamins. The 45-65kDa proteins in the frog and human MPF complex is one of the phosphorylated substrates for the cdc-2 kinase. Thus, MPF now has been broadly defined as a kinase activity appearing in somatic cells undergoing mitosis, in species ranging from yeast to man. It was originally defined as an activity uniquely able to cause progression of meiosis in G_2-arrested frog oocytes. Whereas other kinases, such as a cAMP-dependent kinase, mimic some of the effects of MPF when microinjected into frog oocytes, only MPF purified from M-phase cell extracts can rapidly (less than two hours) trigger germinal vesicle breakdown, chromosome condensation, and spindle formation, accompanied by protein phosphorylation, in cycloheximide blocked frog oocytes. These observations raise the important questions: 1) how is MPF activity regulated post-translationally during the cell cycle, and 2) what is the role of the second protein in the complex in generating the kinase activation which signals the cellular changes needed for division in mitosis? Recent findings from the laboratory of Marc Kirschner indicate that the second protein is a cyclin, and that association of the cyclin protein with the cdc-2 kinase homolog is the critical event which turns on the kinase and drives the cell cycle.

What are the cyclins and how do they exert control over the cell cycle? Cyclin was first described as a 55kDa protein whose abundance oscillated dramatically during the rapid and synchronous divisions of marine invertebrate eggs (clams, sea urchin, and starfish). Cyclins have been identified in many organisms including yeast, frog, and drosophila, and have two forms, A and B, defined on the basis of size, and sequence homology. Cyclin is synthesized during interphase (S phase) and is destroyed as cells exit from mitosis (M phase). This cycle of synthesis and proteolytic degradation for cyclins contrasts with the cdc-2 kinase component of MPF, which can be detected as a peptide at a constant level throughout the cell cycle, but whose kinase activity increases from S phase to M phase, and then decreases again as cells exit from M phase (see figure 1). Evidence that cyclin interacts directly with the cdc-2 kinase component of MPF to control entry into mitosis first came from genetic experiments in yeast. In addition to the yeast cdc-2 gene, a second gene, cdc-13, was shown to interact with cdc-2 in an allele specific fashion, and the two gene products were found to be closely associated with each other in a non-covalently linked complex. The cdc-13 gene product was shown to be a cyclin homolog; deletion of this gene in yeast prevented entry into mitosis and the activation of the cdc-2 kinase activity.

Fig. 1

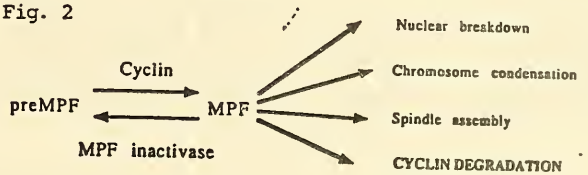


Drs. Andrew Murray and Marc Kirschner set up an in vitro system to define the role of cyclin in MPF activation and cell cycle control. This assay made use of the facts that 1) early embryos have simple cell cycles rapidly moving from S phase to M phase and back to S, 2) in fertilized frog oocytes, protein synthesis in the first part of S phase is required for each mitotic division, 3) the rise in MPF kinase activity with each mitotic cycle required protein synthesis and 4) injection of MPF into protein synthesis-arrested embryos was sufficient to induce all of the events of mitosis. Since it was known that the protein kinase component of MPF remained

constant throughout the cell cycle, and cyclin is synthesized and degraded, Drs. Murray and Kirschner devised a system in which activation of MPF and entry into mitosis depended only on the synthesis of cyclin protein. In an extension of the cell-free system developed by Drs. Maller and Lohka, they made extracts of frog eggs containing nuclei that undergo multiple cell cycles, as assayed by changes in nuclear and chromosomal morphology, DNA replication, oscillations in MPF (and H_1 kinase) activity, and accumulation and degradation of frog cyclin homologs. This cell-free system synthesized proteins prior to M phase, but if protein synthesis was blocked, the nuclei arrested in interphase prior to mitosis. Drs. Murray and Kirschner treated these cell cycle extracts with ribonuclease to degrade all endogenous mRNA, rendering the extracts dependent on exogenously added mRNA for protein synthesis, and causing cell cycle arrest in interphase. When sea urchin or frog cyclin mRNA was added to these RNase-treated extracts, the only protein made was cyclin, which accumulated and was destroyed at each mitosis, demonstrating that cyclin synthesis was necessary and sufficient for entry into mitosis. The rate of the cell cycle was entrained to the rate of cyclin synthesis. Thus, Murray and Kirschner, using their RNA-dependent cell-free system, proved that cyclin is the only newly synthesized protein required to induce MPF activity and mitosis.

This observation also immediately implies a corollary, namely, that cyclin degradation is necessary for exit from mitosis. In a second set of experiments, Drs. Kirschner and Murray went on to demonstrate this point and to propose a model in which cyclin is required for MPF activation, and active MPF directs the degradation of cyclin. The requirement for cyclin degradation for exit from M phase and MPF inactivation was demonstrated by producing mutant cyclins in which the first 13 or 90 amino acids were deleted from the N-terminus. The 13 amino acid deletion was fully capable of driving the cell through multiple division cycles; however, the 90 amino acid deletion mutant only was capable of activating MPF and triggering entry into mitosis. It was not degraded and as a result, produced a metaphase arrest, proving that cyclin degradation was required for exit from M phase. This experiment also reveals that the cyclin sequence needed for MPF activation is distinct from the target sequence required for cyclin degradation, presumably by a protease. Since it is known that MPF phosphorylates its own cyclin subunit, one possible explanation is that the sequence needed for degradation is phosphorylated, and that the active kinase in this way ultimately regulates cyclin degradation.

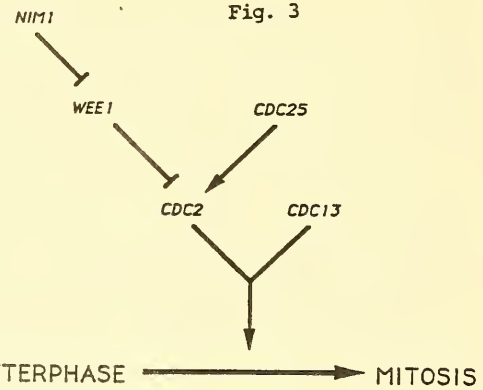
Fig. 2



Thus, studies on regulation of the cell cycle have come a long way. A kinase universally present in germ cells undergoing meiosis, as well as somatic cells undergoing mitosis, is a critical regulator of M phase. Its kinase activity is dependent upon association with a second protein, named cyclin, which is synthesized and degraded with each cell cycle. The association of cyclin with the kinase seems to be necessary and sufficient to drive the cell cycle into M phase, and the degradation of cyclin, leading to MPF inactivation, appears to be required for exit from M phase and thus preparation for future cycles. Once cyclin is synthesized and associated with the cdc-2 kinase in early interphase, there is still a lag before the complex becomes active, indicating further post-translational modification and control. Drs. Kirschner, Murray, Beach, Maller, and others have proposed that the kinase must be phosphorylated in order to associate with the cyclin, after which the kinase becomes dephosphorylated and the cyclin is phosphorylated, leading to conversion of inactive pre-MPF to active MPF. Once activated, histones and other nuclear components are phosphorylated leading to mitosis, and cyclin is degraded leading to inactivation of MPF and return to interphase. While the controlling elements appear to be the kinase/cyclin complex, it is clear that other factors regulate this complex - e.g. a kinase which phosphorylates the cdc-2 kinase, a phosphatase which dephosphorylates the kinase, and a protease which degrades cyclin.

In the frog embryo, which has rapid and relatively simple cell cycles, lacking much additional feed-back control, Drs. Murray and Kirschner postulate at least two additional regulatory factors - an inhibitor (INH) of MPF activity, which is overcome by cyclin, and a cytostatic factor (CSF) which prevents cyclin degradation. In somatic cells there is a far more complex level of feedback control for mitosis. As an example of this, in yeast and mammalian cells a third protein, termed suc-1, associates with the MPF complex and may further regulate activation. Three other genes have been isolated from yeast, wee-1, nim-1, and cdc-25. These gene products and their homologs in higher organisms exert additional positive and negative control (see figure 3). Despite its ups and downs, the field of cell cycle control has come of age, but there is still much to be learned.

Fig. 3



Marion Zatz, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 18, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "A New Address for Lymphocyte Traffic"; GM 37734-03, Butcher, Eugene, Stanford University, Stanford, California

To Director, NIGMS
Through: Program Director, CMBD, NIGMS CAM

Tissue-specific expression of adhesion molecules plays a large role in cell movements in a variety of systems - for example in neural development, tumor cell metastasis, the migration of monocytes and neutrophils to sites of inflammation, and trafficking of lymphocytes from one organ to another via the bloodstream (see Fig. 1.) Extravasation of lymphocytes to organs and tissues involves a very specific cell recognition event, in which a lymphocyte binds via a surface glycoprotein receptor to high endothelial venule (HEV) cells, and then migrates through the HEV to reach the tissue parenchyma (see Fig. 2).

Three independent lymphocyte-HEV recognition systems have been defined thus far which mediate 1) lymphocyte binding to HEV in peripheral lymph nodes, 2) lymphocyte binding to mucosal lymphoid tissues (e.g. Peyer's patches and appendix) and 3) binding to synovium HEV in inflamed joints. In order to understand the molecular basis for these physiologically important tissue - specific cell adhesion events, Dr. Butcher has generated monoclonal antibodies to the endothielal cell molecules which act as specific ligands for the lymphocyte "homing" receptors. (Dr. Butcher has pioneered the identification of these 85-95 kDa lymphocyte receptor molecules as well). Using this panel of monoclonal antibodies, Dr. Butcher has identified and begun to characterize these endothelial cell molecules and study their distribution in both normal and pathologic states. Because the apparent role of these tissue-specific molecules is to convey positional (tissue address) information to circulating lymphocytes, Dr. Butcher has named these molecules "vascular addressins".

Antibodies to both lymph node-specific and mucosal-specific endothelial cell antigens, designated MECA 79 and MECA 367, have been produced. (MECA is an acronym for mouse endothelial cell addressin.) The MECA 367 antibody recognizes a 58-66 kDa glycoprotein selectively expressed on venules of mucosal tissue, whereas the MECA 79 antibody recognizes a 90 kDa glycoprotein on peripheral lymph node blood vessels. Interestingly, mesenteric lymph nodes express both addressins; some endothelial cells stain with MECA 79 antibody, some with MECA 367 antibody, and some have a mottled staining pattern reflecting expression of both molecules. Thus in addition to tissue specificity, there may be micro-regional selective expression as well. Using MECA 367 fluorescently tagged antibody, Dr. Butcher has been able to demonstrate this addressin on both lamina propria and mammary cell vessels, suggesting that this molecule functions as a common recognition element for

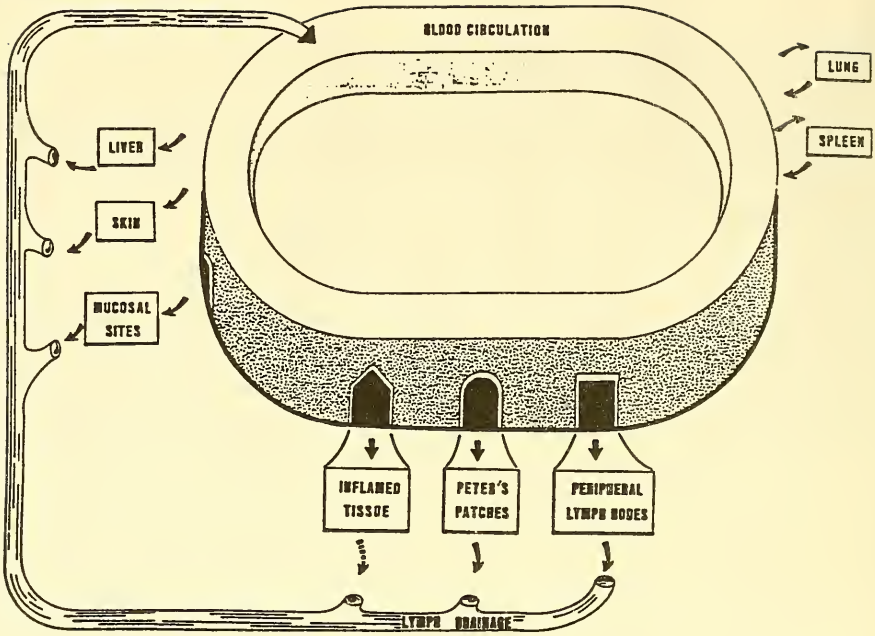


Fig. 1. Illustration of lymphocyte recirculation. Lymphocyte migration pathways are indicated by arrows. Selective migration into tissues such as Peyer's patch, peripheral lymph node and certain inflammatory sites (synovia) is represented by tissue-specific (endothelial) entry ports.

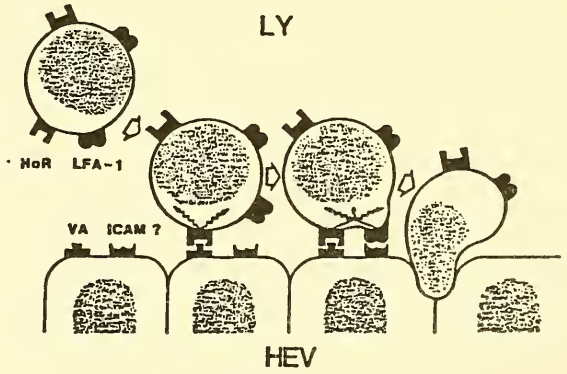


Fig. 2. Speculative view of lymphocyte (Ly) extravasation via high endothelial venules (HEVs), and the involvement of adhesion molecules. HoR, homing receptor; VA, vascular addressin.

lymphocyte homing to diverse mucosal sites, including both organized lymphoid, and extralymphoid tissues. This observation may provide a mechanism to explain the unification of immune responses in widely separated mucosal tissues into a common mucosal immune system.

In further studies of MECA 367, Dr. Butcher has been able to insert the molecule into artificial planar phospholipid membranes and can specifically block the binding of lymphocytes which normally adhere to mucosal HEV by use of the MECA 367 antibody. Both normal lymphocytes and lymphoma cell lines with receptors specific for mucosal HEV bind to these model membranes. Use of this model membrane system should allow further studies of the molecular basis for cell adhesion and provide a better understanding of how and why selective cell positioning or targeting occurs . The lymphocyte-HEV recognition system thus appears to be one of the first models for heterotypic cell recognition for which putative interacting molecules on both cell types have been identified.

The studies on the tissue-specific addressins may prove to be equally important for an understanding of a variety of normal and pathologic conditions. In addition to the broad implications for lymphocyte homing, tumor metastasis, and morphogenesis, Dr. Butcher has demonstrated other systems in which addressin expression is involved in targeted cell-recognition. As examples: lymph node addressin expression has been demonstrated in the subcortical vessels of the thymus and may be involved in the trafficking of precursor T-cells to the thymus; in the NOD mouse model of diabetes, a mucosal addressin is expressed in the pre-diabetic pancreas and may facilitate infiltration of pancreatic islets by autoimmune lymphocytes; in inflamed skin, post-capillary venules have been identified which bind neutrophils (but not lymphocytes or monocytes) thus denoting possibly a third endothelial addressin which mediates inflammatary cell migration.

Marion Zatz

Marion M. Zatz, Ph.D.

Memorandum

Date March 28, 1989

From Program Administrator, CMBD, NIGMS

Subject RESEARCH HIGHLIGHT: "Gene Rearrangements in the Immune System," I. Generation of Diversity; Capra, J. Donald (P01 GM31689-05), University of Texas Health Science Center, Dallas, TX

To

Director, NIGMS
Through: Program Director, CMBD, NIGMS

The immune system is capable of recognizing an almost infinite variety of foreign antigens and generating an immune response (antibodies or cytotoxic lymphocytes). For many years one of the mysteries of immunology has been the question of how so much diversity could be generated, since it was impossible to imagine that an organism could carry in its germ line 10^9 or more antibody genes. A first glance at the protein molecule does little to resolve this dilemma. The antibody protein is a tetramer consisting of two light chains and two heavy chains. The light chains and heavy chains contain a constant region linked to a variable sequence; this variable sequence itself contains three hypervariable domains which differ from one specific antibody to another, and are largely responsible for the molecule's unique interactions with antigen. It was only when scientists began looking at the genes which gave rise to these variable protein molecules that the mechanisms of immune diversity began to be understood.

The variable region of a heavy chain peptide is transcribed from multiple variable region gene segments, termed V-genes (for variable), D-genes (for diversity) and J-genes (for joining). The ability of an antibody molecule, or the related T- or B-lymphocyte receptors, to recognize a diverse array of antigens is based upon the rearrangements of these gene segments (see Fig. 1). At the immunoglobulin heavy chain locus (IGH), five mechanisms are known to operate to generate diversity: 1) Combinatorial diversity by selection of different V_H, D_H, or J_H gene segments to form a complete V_H transcript; 2) junctional diversity generated by variations in the joining of the V_H, D_H and J_H segments,; 3) N-segment diversity, the addition of nucleotides at the junctions of rearranged V-D and D-J gene segments; 4) somatic mutation of these gene segments (operant mostly in secondary immune responses); and 5) differential pairing of heavy and light chains to form different heterodimeric molecules.

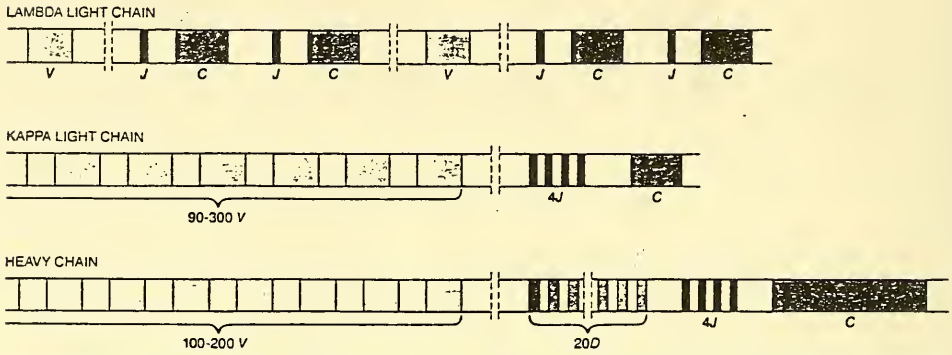
Much is already known about how immunoglobulin and immune receptor genes rearrange. The coding regions of the gene segments are flanked by conserved heptamer and nonomer sequences (see Fig. 2). The heptamer is separated from the nonomer by a non-conserved spacer sequence of 12 or 23 base pairs (bp) (1-2 turns of the helix). This highly conserved sequence motif is found in all species thus far studied. It is presumed that these non-coding sequences

are the binding or recognition sites for enzymes (recombinases) involved in the gene rearrangement process. During the gene rearrangement, the heptamer sequences are precisely joined head to head without any deletion or addition of nucleotides. In contrast the coding sequences are joined in a more random fashion. Ligation of two segments occurs at variable nucleotide positions, resulting at times in significant nucleotide deletions. Alternatively, nucleotides may be added at the joints by the enzyme, terminal deoxynucleotidyl transferase (Tdt). These imprecisions in the joining of gene segments (junctional and N-segment diversity) generate much more somatic variability than if the recombination was precise. This sequence motif is such that a signal sequence with a 12 bp spacer always recombines with a sequence with a 23 bp spacer, generating the "12-23" rule. Thus V-D segments can join, as can D-J segments, but not V-J segments. Presumably D-D segment fusion also would be precluded (see Fig. 2).

Recently however, Dr. Capra has described the fusion of D segments, thus providing an additional mechanism which adds another dimension to IgH variable region diversity. He has shown that, two D_H gene segments can fuse with each other, in either direct or inverted gene rearrangements. These rearrangements have been mediated either by heptamer-nonomer or heptamer-heptamer rearrangementsand also violate the "12 - 23" rule. These fused D segments are fully able to then join with V_H or J_H regions, thus D segments may be utilized in either transcriptional orientation. The resultant V_H-D_H-D_H-J_H recombinations add to the potential repertoire both by increasing combinatorial diversity, and by providing additional sites for N-region variation in the third hypervariable domain of the heavy chain, which is the most critical region for antigen recognition by the antibody molecule.

When one realizes that hundreds of V, D and J gene segments exist, each of which can rearrange in multiple combinations, that the precise joints can vary, as can the combination of heavy and light chains, it becomes apparent that an incredibly flexible system has evolved to accommodate the billions of antigens which an organism can encounter.

Marion M. Zatz, Ph.D.



GENES FOR ANTIBODIES are broken up into small segments scattered widely throughout the genome. Two kinds of light chain appear in mammalian antibody molecules. For the lambda light chain of the mouse there are two *V* genes that encode most of the variable region and four *C* genes for the constant region. Upstream of each *C* gene is a short segment of DNA designated *J*, for joining, which specifies the remainder of the variable region. Either *V* gene can be combined with any pair of *J* and *C* genes. For the kappa light chain there are a few hundred *V* segments, four *J* segments and a single *C* gene. The heavy-chain genes are similar, except that the DNA for the variable region is further subdivided: in addition to the *V* and *J* segments there are about 20 *D* (for diversity) segments. Each set of genes is on a different chromosome. The *T*-cell-receptor genes are organized much as the heavy-chain genes are.

Figure 1

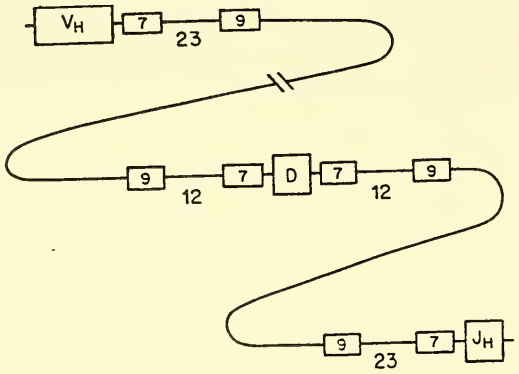


Figure 2

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 7, 1988

From Program Administrator, CMBD

Subject Research Bullet: Metabolism and Function of Complex Sciences
GM 19822-28S1 (Eugene P. Kennedy), Harvard Medical School

To

Director, NIGMS
Through: Director, CMBD

Eugene P. Kennedy and co-workers of Harvard Medical School (GM19822-28S1) have discovered a new function for a long-known protein. This protein, called Acyl Carrier Protein (ACP), previously was known to be essential in the biosynthesis of lipids, by serving as a carrier or handle for the lipid molecule as it underwent a series of biosynthetic reactions. In recent work with the bacterium E. coli, Kennedy and his coworkers have found that ACP also participates in transglucosylation reactions,in which glucosyl residues are transferred from uridine diphosphate glucose to a glucosylated molecule. Transglucosylation reactions are part of the pathway for the synthesis of cell surface carbohydrates which, in turn, probably are important in cell-cell recognition. For example, related work by J.A. Downie and coworkers provides evidence that an ACP - like protein is involved in the synthesis of a carbohydrate which is required for the formation of nitrogen - fixing nodules on the roots of leguminous plants. Further research is required to determine whether ACP is of generalized importance in the synthesis of cell surface carbohydrates, as well as the role of its products in cell-cell recognition or other cellular functions.

M. Janet Newburgh

M. Janet Newburgh, Ph.D.

Memorandum

Date March 29, 1989

From Program Administrator, CMBD Program, NIGMS

Subject RESEARCH HIGHLIGHT: Bullets on Recent Research Contributions in 5130 LM, Membrane Structure and Function

To Director, NIGMS
Through: Director, CMBD Program, NIGMS

Here are a few brief synopses of recent, important research findings in the Membrane Structure and Function Program:

1. **Opioid receptors inhibit Ca^{++} channels in sensory neurons.**
GM38178, Dr. Edwin W. McCleskey, Washington University of St Louis.

Opioid peptides are a class of endogenous neurotransmitter. The opioids bind those sensory neurons whose surface contains a opioid receptor and cause analgesia and euphoria. Morphine and its analogs are pharmacoactive because they mimic this natural action of the opioid peptides. Dr. McCleskey is devoted to describing the molecular basis of the modulation of sensation. He has focused on ion channels, molecular gates which play a pivotal role in the electrical behavior of nervous tissue. He discovered that activated opioid receptors (which are bound to morphine or an opioid peptide) uniformly suppress the activity of both types of Ca^{++} channel in rat sensory neurons. Having identified a clear link between this neurotransmitter and the channels, he will now explore the biochemistry of their interplay.

2. **The proline permease is a Li^+/proline symporter.**
GM39096, Dr. Stanley R. Maloy, University of Illinois at Urbana-Champaign.

Transport proteins display a high degree of specificity in discriminating among the vast array of molecules presented to the cell surface. However, the precise mechanisms of substrate binding and translocation are not yet known for any individual transport protein. Dr. Maloy has studied the proline permease of Salmonella typhimurium, an accessible model system. The energy requirement for proline accumulation is obtained through ion-substrate co-transport, and Dr. Maloy has used a combination of genetic and biochemical techniques to identify and characterize both the amino acid and the Li^+ counter-ion binding sites on the permease molecule. He has described a variety of bacterial proline permease variants which lead to altered sensitivity or resistance to Li^+. Dr. Maloy's work has had a sudden and direct impact on neuropharmacology, where Li^+ is used extensively in treating the manic phase of manic-depressive disorders. He is collaborating on exciting work to

determine the clinical relevance of the proline permease, particularly in predicting treatment failure and providing insights into the mechanism of Li^+ tolerance and toxicity.

3. **Brain ankyrin has a high affinity interaction with the voltage-dependent Na^+ channel in the nodes of Ranvier in myelinated nerve axons.** GM33996, Dr. G. Vann Bennett, Duke University Medical Center.

Eukaryotic cells may be viewed as an assembly of proteins associated in specific ways that determine cell shape, provide organization for the cytoplasm, and determine the position of key regulatory proteins. Studies of the human erythrocyte have provided a useful paradigm to approach the description of these protein interactions, using morphologic and biochemical approaches. Dr. Vann Bennett, along with Dr. Daniel Branton, Harvard University (GM31579 and 39686), has been at the forefront of this field. They have shown that the two major erythrocyte membrane proteins, glycophorin A and band 3 (an anion channel), are associated in a membrane-spanning complex which is linked via band 3 to the protein ankyrin. Ankyrin in turn, is linked to the proteins band 4.1 and spectrin (a large rod-shaped molecule). Together, these five proteins comprise a two-dimensional meshwork that supports the lipid bilayer. Dr. Bennett cloned and sequenced the ankyrin gene and discovered by Northern and Southern blot analysis that it is expressed in non-erythroid tissues. He purified brain ankyrin and found that it binds to brain spectrin, as well as underlying cytoskeletal components. To investigate any possible interaction between brain ankyrin and endogenous anion channels, he collaborated with Dr. Kimon Angelides (NS24606). During myelinated nerve function, action potentials must be propagated in highly specific regions. In these neurons, the voltage-dependent Na^+ channel is clustered and immobilized in specialized sites called the nodes of Ranvier. These nodes occur as constrictions at regularly spaced (1mm) intervals along the fiber; at these sites the myelin sheath is absent and the axon is enclosed only by Schwann cell processes. Drs. Bennett and Angelides, using a combination of immunoblotting, rotational diffusion, fluorescence microscopy, and photobleaching technologies found that purified brain ankyrin specifically bound to the voltage-sensitive Na^+ channel (and no other), both in reconstituted lipid vesicles and in solution. In fractionated rat brains, anti-ankyrin antibodies immunoprecipitated an ankyrin-channel complex. It appears from this work that linkage of the sodium channel to ankyrin has the potential to either stabilize channels following their assembly, or to participate in their initial placement at the nodes of Ranvier. A role for ankyrin in maintaining the specialized distribution of integral membrane proteins had not been appreciated from the earlier erythrocyte studies. Dr. Bennett will now focus on ankyrin isoforms and post-translational modifications as possible bases for molecular sorting.

4. **The human Jk blood type epitope is the urea transporter.**
GM18819, Dr. Robert I. Macey, University of California, Berkeley.

Human erythrocytes are a widely used model system in membrane studies because they are easily accessible, well characterized, simple, lack many metabolic enzyme systems and organelles, and are free of much viscous and confounding nucleic acid. Knowledge of erythrocyte cell surface chemistry includes a wealth of information on blood group antigen systems. Thirty-four such systems are presently listed in McKusick's catalog - ranging from the well known and widely distributed ABO and Rh groups to the rare and exotic Wd-a which has been found in four Hutterites with the surname Waldner. Each system may carry between two and fifteen alleles, and extensive biochemical and genetic mapping data is available in the literature.

Dr. Macey employs electron spin resonance techniques to study membrane transport systems in erythrocytes. He had developed theoretical models, measuring permeability coefficients at various penetrant concentrations by kinetic perturbation methods in order to distinguish between transport phenomena using simple diffusion, simple pores, and simple carriers. He validated his model using diethylene glycol (known to utilize simple diffusion), glucose (known to utilize a simple carrier), and methylurea (known to utilize a simple pore). However, when he applied his techniques to study the transport of urea (which has been controversial in the literature) the kinetic constant was too large to give interpretable results. He therefore decided to search for variant red cells, deficient in urea permeability. He discovered that red cells of the Jk (a^-b^-) blood type were deficient in urea transport and would not lysis in 2 M urea (!). Jk (a^+b^-), Jk (a^-b^+), and Jk (a^+b^+) cells had normal urea permeability. The Jk (a^-b^-) cells were normal in their transport of acetamide - thus settling the controversy over whether amides are transported via the urea system. Dr. Macey's work suggests that the Jk epitope is directly linked to the urea transporter and permits the use of genetic techniques to study the molecular basis of urea transport. Interestingly, the Jk, or Kidd, locus is associated with a deletion in the long arm of chromosome 7. It has been shown to be associated with diabetes mellitus, acanthocytosis (a hemolytic anemia associated with "thorny" appearing erythrocytes), and chronic granulomatous disease.

5. A yeast mitochondrial protein contains funtionally redundant import targeting sequences
GM32703, Dr. Scott D. Emr, California Institute of Technology

Over 90% of mitochondrial proteins are encoded in the cell nucleus, translated on free cytoplasmic ribosomes, and then imported to their correct intramitochondrial location. Gene fusion and deletion analyses have been employed in many laboratories to demonstrate that a transient NH_2-terminal presequence is sufficient to direct a heterologous "passenger" protein into the mitochondrion. Comparison of many such presequences from different mitochondrial proteins of different organisms has found no consensus primary amino acid sequence. Most of the functional presequences are rich in basic and hydroxylated residues and free of acidic amino acids, but no precise structural requirement has yet been defined. Dr. Scott Emr had begun a detailed mutational analysis of the precursor of the β-subunit of the yeast F_1-ATPase. He planned to perform a fine structure analysis of the functional requirements of its import sequence. Using oligonucleotide-directed mutagenesis, he discovered that this protein carries three different, nonoverlapping sequences (at residues 5-12, 16-19, and 28-34) which carried redundant import information. The most distal segment, 28-34, is part of the mature protein and is not removed during translocation. Each of the three sequences could separately direct import. Studies in a yeast rho$^-$ genetic background (with aberrant mitochondria lacking the e$^-$ transport chain in their inner membrane) revealed that the three import sequences would act coordinately to optimize import when growth conditions and the mitochondrial membrane potential were unfavorable. Such distinct, cooperative and redundant sequences may be present to insure that key, essential enzymes such as the ATPase are efficiently targeted, even during stressful conditions. Dr. Emr plans to now search for redundant sequences in other key proteins destined for other cell compartments.

Lore' Anne McNicol

Lore' Anne McNicol, Ph.D.

GENETICS PROGRAM

BIENNIAL REPORT 1988-1989

INTRODUCTION

The goal of the Genetics Program is to acquire knowledge about the fundamental processes and mechanisms of inheritance in health and disease. Research supported by the Genetics Program in past years has laid the groundwork for much of the recent progress in understanding and detecting human genetic diseases and will continue to provide fundamental knowledge needed to develop strategies for treating and, ultimately, preventing genetic diseases. The Genetics Program considers genetic diseases to include not only those conditions with single-gene (Mendelian) modes of inheritance and chromosomal abnormalities, but also commonly occurring conditions, such as atherosclerosis and diabetes, that appear to be caused by a combination of genetic and environmental factors.

In recognition of the fact that basic research is of paramount importance in understanding genetic diseases, the major emphasis of the Program is on fundamental problems in genetics, such as the mechanisms and control of transmission and expression of genetic information. Model organisms, such as bacteria, yeast, fruit flies, corn, worms, and mice are extensively used in this kind of research. The Program is pleased that the 1989 Nobel Prize in Chemistry was awarded to two of its grantees, Drs. Thomas Cech and Sidney Altman, for their independent discovery, in the early 1980's, that RNA can have enzymatic activity. One practical consequence of this discovery is the possibility of using RNA enzymes to block the expression of specific genes in cells, leading to a new way of preventing viral infections. These studies beautifully illustrate how basic research, in this case on bacteria and protozoa, can lead to important advances in medicine.

Research supported by the Genetics Program differs from research in genetics supported by many of the categorical institutes in that it is not targeted to the study of any particular genetic disease. Therefore, the Program takes great pride in having supported the development of new concepts and techniques that were directly responsible for the recent tour de force identification of the gene that causes cystic fibrosis and for the many new diagnostic procedures that enable physicians to detect genetic defects.

Basic research supported by the Program has had an impact on areas far beyond genetic diseases. "Genetic" techniques are now standard in nearly every area of biomedical research and are a mainstay of the biotechnology industry, where they are applied to processes in such varied industries as pharmaceuticals, agriculture, and waste disposal. Most importantly, the concepts and techniques arising

from work supported by the Genetics Program enable investigators to ask ever more sophisticated questions and to perform increasingly detailed analyses.

One area where this was particularly true in the past 2 years is that of mapping and sequencing the genomes of complex organisms. Thanks to many technical and conceptual advances that set the stage for it, a new initiative, known as the Genome Project, was begun in the Genetics Program. This project, whose ultimate goal is to determine the entire nucleotide sequence of the human and other organisms, will greatly facilitate the identification of disease-related genes and will undoubtedly uncover new and unexpected aspects of DNA structure and organization. Beginning in Fiscal Year 1990, the Genome Project will be supported by a new component of NIH, the National Center for Human Genome Research.

In the sections that follow we describe some of the most exciting basic research supported by the Genetics Program in the past 2 years--research that we expect will undergird future biomedical breakthroughs. Research highlights that explain, in more detail, the background and significance of many of these research projects are appended. We have divided the descriptions of the research into four sections, "Structure and Rearrangement of the Genetic Material," "The Mechanics of Gene Expression and Control," "The Genetics of Development," and "Medical Applications of Basic Genetic Research." The Genetics Program also supports two very important resources which are widely used by the genetics community: GenBank and the Human Genetic Mutant Cell Repository; these are also described below.

STRUCTURE AND REARRANGEMENT OF THE GENETIC MATERIAL

Scientific investigations into the nature and stability of DNA, the carrier of hereditary information, began shortly after the rediscovery of Mendel's laws in 1900. The concept of genetic recombination was introduced in 1909, and with the establishment of the chromosomal basis of Mendelian inheritance during the next decade, the possibility that the genetic material is not structurally invariant began to be taken seriously. Definitive evidence for physical exchange between chromosomes, however, did not appear until shortly after the elucidation of the Watson-Crick model for DNA structure in the 1950's. In the subsequent years, numerous studies have demonstrated that genetic rearrangements can occur by a variety of mechanisms that differ, in part, by the degree to which the participating DNA molecules share nucleotide sequence homology. The dynamic nature of DNA structure is also illustrated by the existence of numerous pathways for the repair of chemical alterations that result from exposure to radiation or chemical agents. A striking feature of all of these events is that the integrity of the hereditary information is preserved as efficiently as it is for DNA replication, a process that is essential for transmission of genetic material to succeeding generations.

Recent studies of genetic rearrangements have largely focused on elucidating the mechanisms that allow for the precise cutting and joining of DNA molecules. Considerable progress has been made in analyzing homologous recombination, in which the association between the participating DNA molecules is mediated by base pairing between complementary nucleotide sequences. These studies have resulted in the characterization of recombination enzymes, such as the recA protein of E. coli, that catalyze the pairing of complementary sequences.

Comparable analyses of other types of genetic rearrangements, such as the movement of transposable elements, have been more challenging. Transposable elements are segments of DNA that have the ability to move to different sites in either the same or a different genome. Although these elements are of great interest, the frequency with which they move is so low that it has been difficult to perform mechanistic studies. This problem has now been eliminated with the development of in vitro transposition systems. One Genetics Program grantee has developed an in vitro system for the well-studied bacterial transposon Tn10 which should allow characterization of both the intermediate structures and the enzymes that participate in the transposition process (A1).

Another kind of genetic rearrangement that is of great interest is site-specific recombination, which is important for integration of viral genomes and regulation of gene expression. An exciting development in this area is the discovery that some types of site-specific recombination are mediated by enhancer sequences (A2). These sequences, which resemble the enhancers that affect transcription, can be located at some distance from the crossover site and appear to be needed for correct positioning of the recombining segments.

Genetic recombination is also becoming a powerful tool in modern genetics research. With the advent of sophisticated techniques for manipulating cloned genes, it is now possible to introduce any desired mutation into a DNA segment in vitro. Although subsequent analysis of the resulting mutant phenotype can sometimes also be performed in vitro, it is usually desirable to return the altered gene to its natural, in vivo, environment. To accomplish this, the cloned DNA must be introduced by transformation into the appropriate cell where, by homologous recombination, it can replace the wild type segment. This overall process, which is known as gene targeting, occurs very efficiently in certain organisms, particularly yeast and bacteria. In other organisms, however, problems have occurred at either the level of transformation or during the subsequent integration process. Lack of a good transformation system has been particularly acute in the case of plastids, such as chloroplasts and mitochondria. As a result of a collaborative effort by two investigators supported by the Genetics Program, these organelles can now be transformed by a novel system in which a particle gun is used to shoot DNA-coated tungsten microspheres at target cells (A3). Although the

transformation frequency is low, this "biolistic" system has enabled gene targeting to be performed on the chromosomes of plastids.

Gene targeting in mammalian cells has been frustrated by problems at the recombination step. In contrast to lower organisms, cloned segments of mammalian DNA tend to integrate at random when transformed into recipient cells, making it virtually impossible to detect the rare instances of homologous recombination. This serious technological problem has recently been overcome by development of methods that either enrich for or enhance the detection of homologous recombination events (A4). Improved detection was achieved by using the polymerase chain reaction (PCR) to identify cells in which homologous recombination events have occurred, and enrichment for targeted events was accomplished by clever use of selective genetic markers so that cells that have experienced a random integration event will be killed. These methods are certain to have profound effects on the development of animal models for human disease.

The ability of cellular repair enzymes to respond to environmentally induced genetic damage has been recognized for a number of years. Although most work has focused on the removal of intrastrand pyrimidine dimers that are induced by ultraviolet light, it has become apparent that cells can effectively repair a broad spectrum of chemical lesions. In the case of pyrimidine dimers, a number of enzymes have been implicated in a process that involves detection of the lesion, excision of a short single strand containing the lesion, and then synthesis of a new undamaged strand using the complementary strand as a template. Recently, it was discovered that in mammalian cells this process occurs preferentially in transcriptionally active genes (A5). Moreover, within these genes, dimer removal occurs primarily on the transcribed strand. Although unexpected, the discovery that excision repair and transcription are coupled provides an explanation for the rapid recovery of RNA synthesis that occurs in irradiated cells.

Another unexpected link between DNA repair and other cellular processes is the observation that the yeast rad6 gene encodes a histone-specific ubiquitin-conjugating activity (A6). Ubiquitin is a highly conserved protein that serves as a tag for proteins destined for degradation. The discovery that ubiquitination of histones is catalyzed by the product of a gene involved in DNA repair reveals a new role for ubiquitin. The exact function of ubiquitin in repair is not known although there is considerable indirect evidence that ubiquitinated histones are needed to permit access of repair enzymes to damaged DNA.

Recent studies have also demonstrated that localized structural alterations can occur in chromosomes during normal growth and development. One particularly interesting example concerns the events that occur when chromosome ends, or telomeres, are replicated. Telomere replication has long posed a conceptual

problem since the enzymes that duplicate DNA are unable to replicate the ends of linear chromosomes. Consequently, in the absence of a corrective mechanism, each round of replication would result in shortening of the chromosome. Genetics Program grantees have demonstrated that in protozoa this problem is circumvented by the presence of multiple copies of a repeated six-base sequence at the telomere (A7). The number of repeats is variable and offers, possibly by the formation of terminal hairpin structures, a means for replication without loss of genetic material. A recent discovery by Dr. Blackburn's laboratory is that telomere elongation is catalyzed by a ribonucleoprotein which she has named "telomerase." Since short RNA's serve as primers for DNA replication, Dr. Blackburn is exploring the possibility that the RNA component of telomerase plays a direct role in telomere elongation.

A second example of natural structural variation in chromosomes concerns the existence of variable number tandem repeat sequences, known as VNTR's, that occur throughout the human genome. Unlike the telomeric repeats, the VNTR regions do not vary within one individual, but there is considerable variation among individuals, and new alleles have been observed to arise within families. In addition to providing useful markers for genetic analysis, the existence of VNTR's raises fundamental questions about the organization and dynamics of eukaryotic chromosomes. Duplicated regions have long been thought to arise by unequal recombination during meiosis. However, the creation of new VNTR alleles was shown not to be accompanied by exchange of flanking markers as predicted by unequal crossing over (A8). Although the origin of VNTR's remains unknown, this finding will clearly stimulate new thinking regarding the types of mechanisms that can cause the human genome to rearrange.

Studies of the kinds of rearrangements that can occur in bacterial genomes have yielded the intriguing result that large inversions can only occur in part of the genome. One Genetics Program grantee designed a selection scheme to identify inversion events in Salmonella typhimurium (A9). A systematic survey identified a large region, comprising over one-third of the circular genome, where inversions could not be recovered. The failure to detect inversions in this region is apparently due to constraints that prevent the event from occurring rather than to lethal effects associated with the inverted gene order. While not understood, this effect clearly demonstrates that much remains to be learned about the rearrangements that occur even on relatively simple prokaryotic chromosomes.

THE MECHANICS OF GENE EXPRESSION AND CONTROL

The expression of genes begins with transcription, the process of synthesizing a faithful RNA copy. All cells of an organism contain virtually the same complement of DNA (though there are some interesting DNA changes during differentiation [B1]), and each cell

expresses only a small fraction, about 1 percent, of the genes it carries. Different genes are expressed at different times during development and in differentiated cells of different types. Even in unicellular organisms like bacteria and yeast only a small fraction of the genes are expressed at any time. Most genes that are not expressed are not transcribed, and the initiation of transcription is therefore the most important point at which the array of genes being expressed at any time is controlled.

With the help of powerful genetic tools, gene regulation has been studied for many years in bacteria, and many elegant and sophisticated regulatory mechanisms have been uncovered by Genetics Program grantees. Much work has also been directed toward the more difficult task of understanding gene regulation in higher eukaryotes. While gene regulation in bacteria involves both positive and negative regulators, proteins that bind to DNA and either enhance or inhibit transcription, higher eukaryotes more often use positive regulators. Some are proteins that bind to DNA in the region where transcription starts, usually in combination with several other regulatory proteins, and stimulate the initiation of transcription. Other proteins can exert their effect by binding at enhancer sites some distance from the gene.

The elucidation of such regulatory pathways is a major effort in several laboratories supported by the Genetics Program. It is an important effort because of its relevance to the understanding of human development and diseases, including cancer. For example, one laboratory has been studying an avian viral cancer-causing oncogene that is closely related to a normal cellular gene that is involved in the regulation of gene transcription in response to a hormone (B2). This normal gene behaves very much like a classic bacterial repressor, and the abnormality of the oncogene seems to be due to its lack of responsiveness to the hormone.

Transcription is influenced not only by protein effectors but also by the local structure of the chromosome. One important structural feature is the twist, or supercoil, of the DNA. The work of three Genetics Program grantees (B3) has shown that transcription itself creates local DNA supercoiling, which in turn can affect transcription of nearby genes as well as such processes as genetic recombination.

The pace of progress in understanding gene expression has been accelerated by the use of many different experimental systems, each of which offers unique advantages, and by applying the conclusions drawn from studies on one system to guide hypotheses and experiments in others. Important contributions have come from work on bacteria and their viruses, which continues to provide surprises and advances that stimulate work in other areas, as several highlights illustrate. Starting with Mendel's work on peas and continuing on to the discovery by McClintock of mobile DNA elements in maize, major advances have come from genetic studies on plants. Plants still provide fruitful opportunities for studies of development and gene regulation (B4).

The discovery 10 years ago that genes of higher eukaryotes are often interrupted by non-coding sequences, called introns, was completely unexpected. Scientists are still struggling to understand the significance of this discovery. Whether introns are of any value to the organism, and, if so, whether their importance is physiological or evolutionary, are questions that are still being debated (B5, B6). Recent discoveries (B7, B8) have made introns look more like semi-autonomous elements akin to viruses, or at least to transposons, and have raised fascinating questions about the early evolution of life.

When a gene containing introns is expressed, the introns must be removed from the initial RNA transcript before the RNA can be transported to the cytoplasm and used to direct the synthesis of proteins. Introns are removed from these RNA molecules by a process called splicing, which is usually carried out in the nucleus by an apparatus called the spliceosome, made up of small nuclear ribonucleoprotein particles, or "snurps." The study of this process in yeast and higher eukaryotes is supported by the Genetics Program in several laboratories.

While studying the splicing of an intron in Tetrahymena several years ago, Dr. Thomas Cech, a Genetics Program grantee, made the completely unexpected discovery that the splicing reaction was catalyzed by the intron RNA itself. Other demonstrations of catalytic RNA soon followed. This discovery that RNA can function as an enzyme, a property that was previously thought to be unique to proteins, was of great significance in understanding the origin of life. Based on these discoveries, an ancient "RNA world" has been postulated in which RNA served as both catalyst (for its own synthesis and other reactions) and carrier of genetic information. The RNA world hypothesis provides a way out of the chicken and egg problem presented by the previously debated question of whether DNA or proteins arose first. For his discovery of catalytic RNA Dr. Cech shared the 1989 Nobel Prize in Chemistry with another Genetics Program grantee, Dr. Sidney Altman.

Our new appreciation of the capabilities of RNA has stimulated work on the role of RNA in the ribosome, the organelle that is responsible for translating the RNA message (protein synthesis). Ribosomes are made of both RNA and protein, but are mostly RNA. Earlier work on ribosome function had focused on the proteins, but recent work of several Genetics Program grantees, especially Dr. Harry Noller, has greatly advanced the understanding of the structure and function of ribosomal RNA. The details of the mechanism by which the ribosome, along with tRNA adapters, translates mRNA are still not well understood. However, Dr. Noller and others are now looking to the RNA for the basic mechanism, proposing that the modern ribosome reflects its ancient origin in the RNA world.

Complex as it is, the modern bacterial ribosome has been shown to be capable of self-assembly. Self-assembly has been more difficult

to demonstrate for eukaryotic ribosomes, and the work of Dr. Alexander Varshavsky (B9) has now explained why this is so. Some of the ribosomal proteins find their positions with the help of a protein "chaperon" to which they are attached. Without this chaperon, the mature proteins cannot find their positions and ribosome assembly is blocked.

The mechanism by which amino acids are ordered by the ribosome using aminoacyl-tRNA adapters to recognize the nucleotide sequence of the message RNA in triplets (the genetic code) is understood in principle. For some time the genetic code was believed to be universal, but this uniformity was finally upset by the discovery that mitochondria use a slightly different code. The structure of tRNA, as universal as the code, now appears to be different in some mitochondria as well (B10).

The language of this information transfer is that of nucleic acids, involving specific hydrogen bonded tRNA-mRNA base pairs. The step at which translation really occurs is not on the ribosome but when the amino acid is associated with its tRNA adaptor. The dictionary for this translation is provided by a surprisingly diverse set of enzymes, each of which applies its own rules for recognizing its cognate tRNA. The work of several Genetics Program grantees has made significant progress recently in understanding these rules (B11).

The role of the RNA message in protein synthesis has been easiest to understand in principle: its linear sequence of nucleotides, reflecting the sequence of the gene, specifies the sequence of amino acids. It is not only the one-dimensional sequence of mRNA that contributes to gene expression, however. The folding of mRNA can provide elements of three-dimensional structure that function as regulatory sites (B12) or affect the sequence of the protein product (B13).

The theme that runs through studies of the mechanics of gene expression, as in so many areas of biology, is one of profound universality, affirming the unity of life, punctuated by amazing diversity, asserting its resourcefulness.

THE GENETICS OF DEVELOPMENT

Some of the most fundamental questions in biology are concerned with the genetic regulation of growth and differentiation. Progress in this area of research has been rapid recently, owing in part to the years spent by numerous investigators in establishing the basic parameters of a small number of relatively simple model systems. In addition, technical advances in molecular biology in the last decade have now made it possible to tackle questions in developmental biology that were previously unapproachable.

One of the earliest events in development is the division of a fertilized egg into two daughter cells. Many aspects of this process remain unclear. For example, why does each daughter cell give rise to progeny that follow different developmental pathways, and how is the information for this unequal division stored? Some insight into these questions has been gained recently by investigators studying amphibian embryos. Although most of the RNA's in the fertilized egg are distributed apparently randomly, a few maternal RNA's are distributed unequally along the animal-vegetal axis of the egg. These localized RNA's maintain their positions after the egg divides so that each daughter cell receives a different set of maternal RNA's (C1). These RNA's have been partially characterized, but how they function in development is still a puzzle.

Sex determination of somatic cells is another very fundamental process in nature, and again, mRNA plays a surprising regulatory role. In the case of Drosophila (fruit flies), the ratio of X chromosomes to autosomes in any given cell establishes a regulatory hierarchy of genes that determines the male or female phenotype. What is unusual about this hierarchical pathway is that the gene products are not controlled at the level of transcription, as would have been predicted. Instead, the mRNA's that are expressed by these genes differ because of the way they have been spliced (C2).

One of the most exciting discoveries in developmental biology in the past few years has been the identification of homeoboxes, highly conserved DNA sequences that are involved in pattern formation and segmentation. Homeoboxes have been characterized extensively in Drosophila, where the phenotypes that result from mutations in homeobox-containing genes can be easily studied. Although homeoboxes have been identified in other species, it has only recently been shown that they affect pattern formation in these species also (C3). Also important is the recent finding that homeobox proteins act as transcription factors. They affect the expression of other genes by binding to specific regulatory regions of the DNA (C4).

Although only some genes have binding sites for homeobox proteins, all genes studied thus far do contain regulatory regions to which transcription factors bind. Many of these factors are universal ones and are present in many types of cells. Other transcription factors, however, are highly specialized in that they are only found in certain tissues or at certain developmental stages. The binding of such tissue-specific and stage-specific factors to the regulatory regions of a gene is believed to be the basic mechanism for controlling transcription in a developmentally regulated system. This model has been shown to be true for many different genes, but the regulation of each gene differs enough to make it a unique case (C5). A particularly interesting new dimension has been added recently to this model: that of a "master" regulatory protein that can induce the expression of a whole group of genes during differentiation (C6). The effect of such a "master" switch protein changes a cell from one phenotype to another.

During the past 2 years a great deal of progress has been made toward understanding a basic process of all growing cells, the cell cycle. Investigators supported by both the Genetics Program and the Cellular and Molecular Basis of Disease Program have identified many of the proteins involved in this process. It is clear that these proteins are highly conserved in organisms as diverse as yeast and man (C7). There is now good evidence that phosphorylation plays an important role in regulating the timing of cell division (C8). Tied to this process is the cell-cycle control of other genes; certain genes are "timed" to be expressed at specific stages of the cell cycle. This type of regulation works at the level of transcription, and the regulatory factors are abundant only at certain times in the cell cycle, thereby causing induction only at the specified times (C9).

The developing organism responds to both internal and external cues. The genetic "program" that specifies the differentiation of any given cell is influenced by the rest of the organism and by the world beyond the organism. Just as a differentiating cell responds to internal cues to express a certain phenotype, it must also have a way to respond to external cues, such as changes in the chemical or physical environment, stress, or the passage of time. The heat shock response is a well studied genetic response that may function universally to protect cells from any extreme in the external environment. Recent research provides new clues to the mechanisms by which cells respond to heat shock (C10 and C11). Cells also respond to the external change from day to night by establishing a circadian rhythm. This biological timing mechanism appears to have a genetic component which is linked to a proteoglycan that affects communication between cells (C12).

On a broader level, speciation can be viewed as a genetic response to external pressure. The process of natural selection is the ultimate survival mechanism. Subtle variations between organisms become exaggerated as one variant assumes an advantage over another. Hence, a species of fruit fly that can metabolize a certain chemical more efficiently than another establishes a "niche" that eliminates other, less efficient species of flies (C13).

MEDICAL APPLICATIONS OF BASIC GENETIC RESEARCH

One of the goals of the Genetics Program is to support fundamental research that will lead to an understanding of inherited disorders. Although support is not provided for research on specific diseases, many grantees studying basic genetic mechanisms have discovered information that will have a significant impact on clinical research. For example, studies on mitochondrial DNA have had important implications for development of therapies for AIDS (D1). Dr. Melvin Simpson has shown that the toxic side effects of the drug AZT, which is a potent inhibitor of HIV-1 replication, may be due to its inhibition of mitochondrial DNA synthesis. This

observation prompted him to test the effect of other anti-HIV-1 drugs on rat mitochondrial DNA replication. Several which were found to be free of effects on mitochondria are now being tested as less toxic candidates for the treatment of AIDS.

Basic research in lower eukaryotes can also make important contributions to human genetics. Recent studies with Drosophila have suggested a molecular basis for the fragile sites that occur in human chromosomes and that are associated with tumorigenesis and inherited mental retardation (D2). Dr. Charles Laird has found that the comparable sites in Drosophila chromosomes undergo delayed replication and, as a result, appear to undergo abnormal chromosome condensation. These observations have been used to formulate a model that, while still unproven, offers a detailed explanation for the observed patterns of inheritance of fragile chromosome sites. In a similar manner, studies on Drosophila that are unable to undergo normal development have provided insight into the metastatic potential of human cells (D3). By searching the DNA sequences in GenBank, a Genetics Program grantee discovered that the awd gene in Drosophila, which is required during the second half of larval development, shares extensive homology with a human gene, nm23, whose loss is correlated with high metastatic potential. Moreover, the mutant phenotypes of Drosophila and human cells lacking, respectively, functional awd and nm23 genes are strikingly similar. Analysis of the awd gene in fruit flies may thus provide insights into the development of cancer in humans.

Modern genetic analysis has also made contributions to our undertanding of the origin of single-gene disorders. Dr. Alan Scott has recently discovered that hemophilia can arise by insertion of large DNA sequences into the gene that encodes Factor VIII of the human blood clotting system (D4). The insertions were found to be members of a family of repeated sequences that are believed to undergo transposition by a mechanism that involves an RNA intermediate and reverse transcriptase. This process, which is called mutation by retrotransposition, had not been previously implicated in a human genetic disorder.

Finally, a Genetics Program grantee with a longstanding interest in establishing evolutionary relationships among organisms based on their of ribosomal RNA sequences found that the organism Pneumocystis carinii is a close relative of yeasts and is evolutionarily far removed from protozoa (D5). The classification of P. carinii, which causes pneumonia in AIDS patients, had until now been uncertain. Determining its biological classification will make it possible to design more rational drugs to treat the infection it causes.

GENBANK, THE NUCLEOTIDE SEQUENCE DATA BANK

GenBank has just completed its seventh year of operation. The database was initiated as a pilot project in 1979 by Dr. Walter Goad at Los Alamos National Laboratory (LANL). In 1982, a 5-year

contract was awarded by the National Institute of General Medical Sciences to Bolt, Beranek and Newman, Inc. (BBN) of Cambridge, Massachusetts, working in collaboration with LANL, to establish GenBank. The renewal was awarded in 1987 to Intelligenetics, Inc. of Mountain View, California, with a subcontract for data collection again to LANL. Throughout its history, GenBank has coordinated the task of data entry with the European Molecular Biology Laboratory (EMBL) Data Library. This collaboration has recently been extended to include the DNA Databank of Japan (DDBJ). In addition, with the cooperation of the U.S. Patent and Trademark Office, sequences contained in patents will be included. The data are comprehensive and available to all scientists throughout the world. The importance of sequence data is best illustrated by the recent initiative to map and sequence the human genome where it is felt that the sequence will have a profound impact on our understanding of the causes and possible treatments and cures of genetic diseases.

As of September, 1989, GenBank had made 61 releases of sequence information totaling 28,791 entries (corresponding to 34,762,585 base pairs). GenBank data are released quarterly on magnetic tape and floppy diskette. The most recent data are available through an on-line service, which now operates under a user-friendly interface, IRX, developed by the Lister Hill Center of the National Library of Medicine. The typical GenBank entry contains the nucleic acid sequence, a table that lists important features of the sequence, other biological information, such as the organism from which the sequence was derived, and the literature citation. By far the largest number of users access GenBank through secondary distributors such as The Protein Identification Resource, Prophet, or commercial packages that include analysis software.

GenBank is an interagency project with NIGMS taking the lead role of working directly with the contractor and coordinating the funding. Drs. James Cassatt and Irene Eckstrand are the Project Officers. The total budget for the current 5-year contract is about $19,000,000. Sponsors include, in addition to NIGMS, the National Center for Human Genome Research, the National Library of Medicine, and the Department of Energy. The Department of Energy also provides support for research related to database design at Los Alamos.

At the time of the last report 2 years ago, GenBank contained about one-half of the data it contains now. This means that it took 5 years to accumulate the first 14,000 entries and 2 years to accumulate the next 14,000. To accommodate the ever-increasing rate of data production and in anticipation of an even greater rate as a result of the human genome initiative, GenBank and its collaborators have reached agreement with most of the journals that publish significant amounts of sequence data--an agreement that makes publication of a paper contingent on direct submission by the author to GenBank of the sequence data contained in the paper. To streamline operations at GenBank, the data are being installed in a transaction-driven relational database management system. This

system will make it possible automatically to update the data which arrive from collaborating databases. In addition, GenBank has recently completed and is now distributing a PC-based computer program that can be used by scientists to submit their data directly to GenBank. An essential feature of this program is that the output is in the form of transactions that can be automatically entered into the database. The implementation of these procedures will ensure that GenBank will remain current with the scientific literature.

Finally, the Project Officers together with staff from the National Library of Medicine and the National Center for Human Genome Research are beginning the planning process for the next generation of GenBank which begins with the expiration of the current contract in September, 1992.

THE HUMAN GENETIC MUTANT CELL REPOSITORY

The Human Genetic Mutant Cell Repository, currently in its 17th year of operation, continues to serve the genetics and cell biology research communities by providing well-characterized, contamination-free cell lines from patients with a wide range of genetic disorders and from unaffected family members. The Repository, also known as the Cell Bank, is supported by a contract which is recompeted every 5 years. On January 1, 1989, a new contract was awarded to the Coriell Institute for Medical Research in Camden, NJ, the organization which has housed the Repository since 1972.

The Repository has been constantly evolving to meet the changing needs of researchers and to anticipate their future needs. It continues to acquire samples representing well-characterized biochemical and cytogenetic defects. In response to the great opportunities for mapping the locations of disease-related genes, the Repository also devoted considerable effort over the past several years to acquire cell samples from families with genetic disorders whose causes are not yet understood. Cell lines established from these samples are used to locate DNA markers for the disease and ultimately to identify the disease-causing gene itself. Cell lines from multigenerational families with retinitis pigmentosa, Usher syndrome, von Recklinghausen neurofibromatosis, and tuberous sclerosis have recently been added to enhance the collection of cell lines from other multigenerational families, including those with cystic fibrosis, maturity onset diabetes of the young, and ataxia-telangiectasia.

The Repository also maintains cell lines from extended pedigrees of normal individuals (Utah collection) and of individuals with primary affective disorders (Old Order Amish collection) and Huntington disease, which are of great value for gene linkage studies, and an extensive collection of chromosomal deletions, duplications, and translocations, which are used to map genes.

The Repository has recently initiated a major new collection--somatic cell hybrids--which will provide an important resource to researchers interested in physical mapping of genes. Somatic cell hybrids are frequently unstable in culture, leading over time to loss of human chromosomes. To assure that researchers receive material with the expected human chromosomal complement, the Repository will not, in general, supply the cell cultures themselves. Rather, they will send DNA samples isolated from the somatic cell hybrids at the same passage at which they were characterized by Southern blots and in situ hybridization.

The collection currently contains over 4,600 cell lines representing more than 350 distinct genetic disorders. The Repository ships approximately 4,000 samples each year to researchers in the United States and abroad. Among the cell lines requested most frequently in the past year were those from the Utah collection and the Old Order Amish collection, as well as those from patients with diabetes mellitus, xeroderma pigmentosum, Huntington disease, and chromosomal abnormalities, and from normal individuals.

The Repository publishes a catalog or supplement annually. The catalog serves as a valuable reference to investigators in genetics and cell biology. In addition to listing the holdings of the Repository, it contains diagrams of the cultures available with translocation breakpoints and regions of monosomy, trisomy,or greater dosage for each human chromosome. Cell lines are cross-referenced by their McKusick numbers. This complements the listing of the Repository's holdings which is included in Victor McKusick's Mendelian Inheritance in Man.

The project officer of the Repository is Dr. Judith Greenberg. She is advised by the newly established Human Genetic Mutant Cell Repository Working Group, whose members serve 4-year terms. The Working Group meets at least twice a year to evaluate the progress of the Repository, to recommend future acquisitions, and to help set policy. In addition, members serve as curators, reviewing all newly submitted samples before being made available to the community.

ORGANIZATION AND STAFFING

The Genetics Program employs a full range of mechanisms to support research and research training in genetics. These include regular research grants (R01), FIRST Awards (R29), program projects (P01), centers (P50), Academic Research Enhancment Awards (R15), Small Business Innovative Research grants (R43, R44), institutional predoctoral and postdoctoral training grants (T32), individual postdoctoral fellowships and senior fellowships (F32, F33), and contracts.

The Genetics Program, directed by Dr. Judith Greenberg, is divided into two sections, the Molecular and Medical Genetics Section and

the Genetics of Growth and Differentiation Section. The Molecular and Medical Genetics Section supports studies on the replication, recombination, and repair of DNA; the mechanisms of mutagenesis; extrachromosomal inheritance; protein synthesis; human medical genetics; and gene mapping. The head of this section is Dr. Irene Eckstrand. She is assisted by two program administrators who have joined the Program within the last 2 years, Dr. Marcus Rhoades and Dr. Stephen Fahnestock. Dr. Mark Guyer, formerly in this section, left to join the National Center for Human Genome Research.

The Genetics of Growth and Differentiation Section supports studies on chromosome organization and mechanics; developmental genetics; mechanisms and control of gene expression; rearrangement of genetic elements; neural and behavioral genetics; and population genetics. Dr. Judith Greenberg is the acting head of this section. She is assisted by three program administrators, Dr. Barbara Williams, Dr. Joye Jones, and Dr. Catherine Lewis. Dr. Lewis has recently joined the Program. Dr. Jane Peterson, formerly in this section, left to join the National Center for Human Genome Research.

APPENDIX

RESEARCH HIGHLIGHTS

A. STRUCTURE AND REARRANGEMENT OF THE GENETIC MATERIAL

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 17, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "In Vitro Transposition"
Nancy E. Kleckner, Harvard University (R01 GM 25326)

To Director, NIGMS
Through: Acting Director, Genetics Program, NIGMS 3-18-88

Transposable elements are DNA sequences that are capable of moving to many different sites in the genomes of bacterial, plant, and animal cells. Their movements result in genetic rearrangements, which may have pronounced effects upon gene expression in the cell. This type of transposition and genetic rearrangement occurs outside of the usual pathways known for recombination. Although many models have been proposed, very little is known about the mechanism of transposition. Most of the information on transposition has come from the study of bacterial transposable elements, because the genetic systems in bacteria are better defined than those of higher organisms. The study of bacterial transposable elements has been important not only because they serve as a model for understanding genomic rearrangement in higher organisms, but also because bacterial transposable elements carry genes that confer resistance to antibiotics.

Dr. Nancy Kleckner and her associates have been studying a transposable element known as Tn10, which is found in bacteria and carries the gene for tetracycline resistance (1). Their work has helped to elucidate the DNA structure of Tn10 and many functional aspects of its component elements. Most of this work was based on in vivo analyses, but more detailed studies of the biology of Tn10 required the development of an in vitro transposition system. Many scientists have attempted to develop such a system, but most have met with little success. Dr. Kleckner has now succeeded in developing an in vitro transposition system for Tn10.

The structure of Tn10 is illustrated in figure 1. The element, which is 9300 base pairs (bp) in length, carries the tetracycline resistance gene flanked on each end by a 1400 bp repeated sequence. These inverted repeat ends are called IS10-right and IS10-left to indicate their relative positions on the element and their homology to smaller bacterial transposable elements called insertion (IS) elements. The IS10 ends encode the sites and functions needed for Tn10 transposition, and they cooperate with one another to mediate transposition of the tetracycline resistance genes. IS10 has never been found in association with any transposable element other than Tn10. Transposase, an enzyme required for transposition, is encoded by IS10-right; IS10-left is functionally defective. Further dissection of the IS10 elements demonstrated that the "outside" and "inside" termini of IS10 (defined with respect to their position within intact Tn10),

are active sites for Tn10-promoted DNA rearrangements. Studies of DNA methylation within Tn10 revealed that there are two DNA adenine methylation sites (called dam sites) that regulate transposition.

The development of an in vitro system for functional studies of Tn10 was stimulated by the discovery of double stranded circles formed as an end product of transposition (2). These circles are thought to be the result of double stranded excision of the element from the host DNA prior to replication. Dr. Kleckner believes that these circles are not intermediates in the transposition process, but rather result from Tn10 DNA that is successfully excised but which fails to find host DNA in which to insert.

The in vitro system designed by Dr. Kleckner's group involves the incubation of a donor Tn10 transposable element carrying ampicillin resistance and target lambda phage DNA in a crude E. coli extract lacking an added energy source. In bacteria treated with Tn10, phenotypic transposition based on resistance of cells to ampicillin was observed at a very low level (10^{-6} events per donor plasmid molecule). However, transposable element circles identical to those observed in vivo were detected at a level of 10^{-2}. The presence of circles was easily detected on Southern blots in which the DNA from the reaction was probed with Tn10 sequences. The circle assay provided an easy and reliable method to measure transposition in vitro.

The next step was to identify the active factors in the system by altering the components of the extract. When a purified fraction of transposase was incubated with a substrate containing only outside end IS10, no circle formation was observed. However, the addition of a protein-containing extract from an E. coli strain that lacked transposase restored circle formation. These results suggested that, in addition to transposase, other E. coli proteins were required for transposition to occur. Fractionation of the crude extract demonstrated that factors belonging to a class of small, basic proteins, previously implicated in several types of recombination and replication reactions, were required for transposition as measured by circle formation.

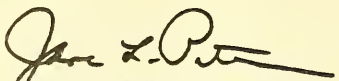
One of these factors is a sequence-specific DNA binding protein called Integration Host Factor (IHF). Its binding to DNA is required for site-specific recombination. The role of IHF in IS10 transposition was assayed in experiments where a crude extract from IHF^- cells (lacking IHF) was used. The formation of circles using this extract was reduced compared to that observed in the IHF^+ extract. The addition of purified IHF to the IHF^- extract stimulated additional circle formation, confirming the direct participation of IHF in transposition. However, the presence of a low level of circle-forming activity in the IHF^- extracts suggested that another host factor might also be active in promoting circle formation. Purification of an additional active fraction from the IHF^- extract revealed that the active protein was a prokaryotic "histone-like" protein called HU. HU is related to IHF in primary structure but, unlike IHF, it shows no binding specificity. Further experiments showed that HU could act alone as a host factor in promoting transposition, albeit with a lower specific activity than IHF.

The presence of a third stimulatory factor (Factor Y) was detected in other fractions resulting from the HU purification. Factor Y does not directly supplement transposase activity but acts by enhancing the effects of IHF and HU. Limited biochemical characterization demonstrated that it enhances the activity of the other factors and is not sensitive to protease digestion. Quantitatively, the combined activity of the IHF, HU, and Factor Y account for the total activity in the crude extract.

In order to determine the DNA sequence requirements for transposition, Dr. Kleckner studied the effects of changes in the IHF binding site found near the outside end of IS10. As expected, when DNA lacking the binding site was used in the system, no circle formation occurred. Surprisingly, when two IS10 inside ends were present on the plasmid, circle formation proceeded efficiently in the presence of transposase alone. IHF and HU had only a very slight stimulatory effect on such a reaction. In a reaction in which the DNA containing one inside and one outside end was used, circle formation was stimulated by the host factors to the same extent as for DNA containing two outside ends.

Further analysis of the outside and inside ends of IS10 is needed to understand why circle formation can proceed with DNA containing only the inside ends in the absence of host factors. There are two possible explanations. The outside ends may have evolved so that they are functional only in the presence of host factors. Alternatively, the outside ends may have acquired inhibitory sequences, whose effects are overcome by host proteins. Dr. Kleckner believes that the outside end may serve a regulatory function for the host cell. For example, if the level of IHF varies depending upon growth conditions, transposition frequency, mediated by IHF binding to the outside end, would vary accordingly. Interestingly, in its other roles in the cell, IHF is involved in such regulation of enzymatic activity.

The development of an *in vitro* system for Tn10 transposition offers an opportunity to answer a new set of questions about transposition. For example, how and where do the host factors act during circle formation? What is the actual mechanism of Tn10 transposition? Can other elements transpose in this *in vitro* system and, if not, what additional factors are required? What is the energy source for transposition? As in the development of the system, these studies are likely to yield not only new insights about the process itself but also information about recombination and host/plasmid interactions.



Jane L. Peterson, Ph.D.

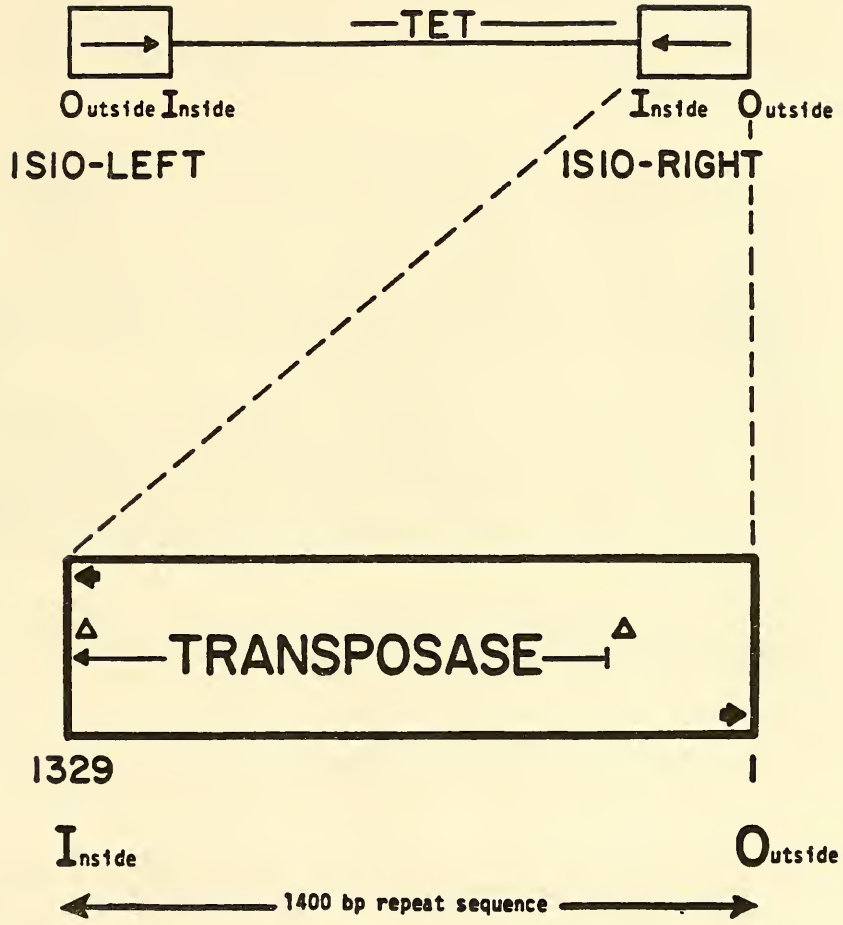
References

1. Kleckner, N., Morisato, D., Roberts, D., and Bender, J. (1984). Mechanism and Regulation of Tn10 Transposition. *Cold Spring Harbor Symposia in Quantitative Biology*, vol. 49, pp. 235-249.

2. Morisato, D. and Kleckner, N. (1988). Tn10 Transposition and Circle Formation *In Vitro*. *Cell*.

Figure 1

Tn10



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 26, 1988

From Chief, Molecular and Medical Section, Genetics Program, NIGMS

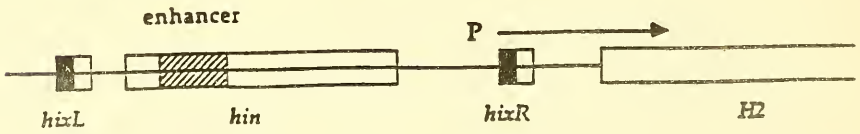
Subject: RESEARCH HIGHLIGHT: "How to Get It Backwards or Hin Site is Better Than No Site at All" GM 38509 (Reid Johnson and Mel Simon), University of California

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 7-26-88

Only a few years ago, scientists discovered that there are special DNA sequences that enhance the transcription of specific genes even though the gene and the enhancer sequence may be very far apart on the DNA strand. Enhancer sequences are now known to be binding sites for specific nuclear proteins; however, nobody knows exactly how they activate transcription. Recently, enhancer sequences which stimulate not transcription but DNA recombination have been discovered in prokaryotes. Because recombination enhancers can be studied in simple in vitro systems, they may be excellent models for studies of other enhancer sequences.

Site-specific recombination in bacteria is the basis of integration or excision of a plasmid or bacteriophage into the bacterial genome or, in some cases, the inversion of a DNA segment. Recombination reactions are carried out by specific proteins called recombinases, which catalyze the exchange of DNA strands at short homologous sequences. Details of the recombination reaction depend on the specific homologous sequences and whether the sequences are directly repeated or inverted, as well as on the specific recombinases involved.

Two NIGMS grantees, Dr. Reid Johnson and Dr. Mel Simon, have been collaborating on studies of a specific recombination system, Hin, in Salmonella. A diagram of the DNA region is shown below:



Drs. Johnson and Simon have found that the complete Hin recombination system requires three proteins and three DNA sequence elements, described below:

Proteins

- Hin - Recombination protein or invertase which binds to the recombination sites, hixL and hixR
- Fis - Factor for inversion stimulation which binds to the enhancer sequence
- HU - Histone-like protein which binds nonspecifically to double stranded DNA

DNA sequence elements

- hixL - The left end of the recombination site
- hixR - The right end of the recombination site
- enhancer - Specific sequence required for high levels of inversion
- H2 promoter (P) - Sequence which activates transcription of the H2 gene

Hin protein regulates the expression of flagellin genes (the H antigen) and is encoded by the hin gene (shown in the above illustration), which lies next to the H2 promoter. Inversion of a 996 base pair (bp) region of DNA between hixL and hixR and containing the hin gene inactivates the H2 promoter and turns off transcription of the H2 antigen gene. If the 996 (bp) region is again inverted, its normal orientation to the promoter is restored, and transcription can proceed. The Hin protein itself is responsible for making double stranded cuts at the recombination sites and is essential for inversion to occur. Because Hin protein is a recombinase that mediates an inversion event, it is called a DNA invertase. Other DNA invertases with similar properties and functions (Pin, Cin, and Gin, for example) are known. Although they are found in different organisms, Pin, Cin, Gin, and Hin proteins share 60-70% amino acid homology and can replace each other functionally.

The enhancer (a 60 bp sequence) is usually embedded in the hin gene but can be moved just about anywhere as long as it remains on the same DNA molecule as the recombination sites, hixL and hixR. Even at a distance of 4000 bp from its usual site, the enhancer functions efficiently. The exception is that the enhancer cannot stimulate recombination when it is too close - within 30 bp of hixL or hixR - although even then it can bind Fis protein quite well. Based on electrophoretic mobility studies, Drs. Johnson and Simon have concluded that the enhancer sequence probably contains bent DNA, which suggests that the structure of DNA in the enhancer region, as well as its sequence, has some significance.

A great deal of work has gone into characterizing each of the recombination proteins, Hin, Fis, and HU. Their sizes, kinetics, and binding sites are known. In addition, the sequences of the DNA elements - hixL, hixR, the

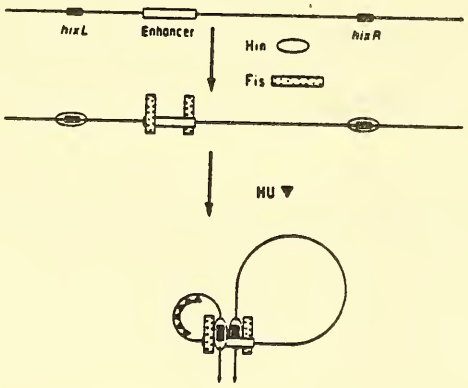
promoter, and the enhancer - have been determined and analyzed in great detail. With this information Drs. Johnson and Simon began to speculate on the nature and function of the recombination enhancer. Any model they developed had to incorporate the following characteristics:

1. The enhancer functions efficiently up to 4000 bp away from either recombination site and in either orientation. It cannot function within 30 bp of the recombination sites.

2. The enhancer has two domains, each of which binds a Fis protein molecule. The integrity and spatial relationship between the domains are critical. There is no evidence for interaction between Fis proteins bound at the enhancer domains.

3. Fis and Hin bind independently to their DNA sequence elements. Hin can make the correct double stranded cuts in the DNA recombination sites even when the enhancer sequence is absent. However, the inversion step (strand exchange) requires the enhancer sequence and all the other components of the system.

The Johnson and Simon model, called the 'looping-synapsis' model, is illustrated below:



In this model, two Hin proteins bind to hixL and hixR; two Fis proteins bind to each end of the enhancer sequence at specific sequence domains. Subsequently, protein-protein interactions between Hin and Fis molecules generate the looped structure shown in the illustration. HU proteins presumably coat the looped DNA and hold it in a specific configuration to facilitate site-specific recombination. The model suggests that the function of the enhancer is to position the recombination sites with bound Hin molecules in exactly the correct configuration to generate synapsis and strand exchange at hixL and hixR.

The Hin recombination enhancer has much in common with transcription enhancers. For example, both transcription and recombination enhancers are often binding sites for proteins which presumably bring specific DNA sites into a specific configuration. As in the case of the Hin system just described, the formation and stabilization of such structures provides an exquisitely precise mechanism for gene regulation.

Understanding the Hin system may also provide insight into the mechanism of site-specific recombination, which requires that specific DNA locations which are targeted for recombination be brought into proximity. This complex and precise process is not well understood, although it is fundamental to many biomedical issues from the integration or excision of transposable elements to gene replacement. The research reported by Drs. Johnson and Simon provides important clues for examining the nature and regulation of recombination.

Irene Anne Eckstrand

Irene Anne Eckstrand, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date June 27, 1988

From Program Administrator, Genetics Program, NIGMS

Subject Research Highlight: "Targeting Plastid Transformation" (Ronald Butow, University of Texas Health Science Center, GM22525; John Boynton, Duke University, GM19427)

To Director, NIGMS
Through: Director, Genetics Program, 6-27-88

Transformation is the term used by geneticists to describe a process in which a cell takes up added, purified DNA and then incorporates the DNA into its genome in a way that allows the information encoded in it to be inherited by that cell's progeny. Transformation has been the methodological basis for many fundamental discoveries in biology, the most important being the identification of DNA as the genetic material. Transformation is also one of the experimental foundations for recombinant DNA technology. The mechanism of transformation differs from one organism to another, but scientists have found experimental conditions that allow many types of bacterial, animal, and plant cells to be transformed and the determination of the specific conditions which allow it to be transformed is often a crucial first step in the use of a particular organism or cell type in a study. Therefore, investigators interested in studying the sub-cellular organelles known as plastids have been frustrated for many years because of their inability to develop a reproducible plastid transformation system.

Plastids serve as the sites of specialized metabolic activities. Two important and well-studied plastids are mitochondria, the sites of oxidative phosphorylation, and chloroplasts, the sites of photosynthesis. Each of these organelles contains DNA that encodes a number of gene products that play essential roles in plastid function. To study the roles of the individual gene products, investigators would like to be able to experimentally manipulate the plastid genes, reintroduce the altered versions back into the plastid DNA, and observe the effect of the alteration in vivo. Until now such experiments could not be done because of the lack of successful plastid transformation systems. However, work in the laboratories of Drs. John Boynton (GM19427) and Ronald Butow (GM22525) has recently led to the successful transformation of both chloroplasts and mitochondria.

In each case, the success came through the use of the so-called "biolistic transformation" system, originally developed by Dr. John Sanford of Cornell University for the transformation of plant cells (1). The transformation system consists of a particle gun (Figure) that uses a gunpowder charge to shoot tungsten microspheres at target cells. The microspheres are

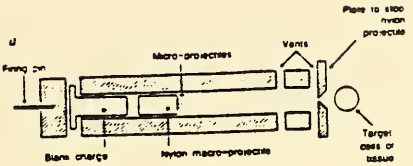


Fig. 1 a, Schematic diagram of the particle gun. About 0.05 mg of tungsten particles (average diameter 4 μm, General Electric Corp., Cleveland, Ohio) is placed on the front surface of a cylindrical nylon projectile (diameter, 5 mm; length, 5 mm) as a suspension in 1–2 μl of water. A gun powder charge (GY22AC, gray extra light, No. 1, Speed Fasteners Inc., St Louis, Missouri), detonated with a firing pin, is used to accelerate the nylon projectile down the barrel of the device. The tungsten particles continue toward the target cells through a small 1-mm aperture in a steel plate designed to stop the nylon projectile. The tungsten microprojectiles leave the particle gun with an initial velocity of about 430 m s^{-1}. This value was determined by allowing the nylon macroprojectile to leave the barrel of the device and estimating its velocity in-flight with a chronograph (Ohler Research, Austin, Texas). The target cells are placed ~10–15 cm from the end of the device.

(figure taken from reference 1)

approximately one to four microns in diameter, and cells can survive being penetrated by as many as several. For transformation, the microspheres are treated with the DNA of interest, which adsorbs to the sphere surface. After the DNA-coated sphere enters a cell, the DNA apparently desorbs from the sphere and interacts with cellular DNA. For example, Dr. Sanford showed that the biolistic transformation system can deliver functional tobacco mosaic virus RNA to onion epidermal cells (1). In collaboration with Dr. Sanford, Drs. Boynton and Butow have shown that transformation of Chlamydomonas reinhardtii chloroplast DNA (2) and Saccharomyces cerevisiae mitochondrial DNA (3) can be achieved with this system. It appears that this mechanical delivery of DNA into the plastid circumvents one of the major problems that had previously stymied plastid transformation, the inability of the added DNA to pass across the membrane that surrounds the plastid.

The experimental plans were comparable for each of the plastid systems. As a target (recipient), a strain containing mutant plastid DNA was used. The mutation was a deletion of a known plastid gene. Use of a deletion mutation achieved two major purposes. It prevented the appearance of normal cells in the experiment by reversion of the mutation. It also provided a unique marker for the recipient DNA, in the form of a restriction fragment whose size differed from the wild type restriction fragment by the size of the deletion mutation. The DNA to be introduced by transformation, or donor DNA, consisted of a plasmid carrying a wild type copy of the mutant gene in the case of the Chlamydomonas experiments. In the case of the yeast experiments, the donor plasmid caried a gene with a deletion that was different from the deletion in

the recipient DNA and which did not interfere with the wild type function of the gene.

After bombardment of the population of mutant target cells with the donor DNA-coated microspheres, transformed yeast or algal cells were found that displayed the phenotype determined by the functional donor plastid gene. No transformants were observed if the microspheres had not been treated with DNA or had been treated with control DNA that did not contain the functional gene. Both biochemical characterization of the relevant gene product and molecular genetic characterization of the plastid DNA in the transformed cells demonstrated that these genetic variants had resulted from the introduction of exogenous plastid DNA, followed by replacement of the mutant recipient DNA sequence with the wild type donor DNA sequence, i.e. that the recipient plastids had been transformed.

In each case, the efficiency of the plastid transformation process was very low compared with other transformation systems. With Chlamydomonas, the frequency of chloroplast transformation with the biolistic method was reported to be ten to twenty transformed cells per microgram of donor DNA applied. In contrast, the best bacterial transformation systems can give between 10^6 and 10^8 transformants per microgram of donor DNA. With Saccharomyces, the frequency of mitochondrial transformation was about 0.1% of the frequency with which a nuclear gene was transformed using the biolistic technique. Current work in both Dr. Boynton's and Dr. Butow's laboratories is devoted to improving the efficiency of plastid transformation, as well as to using the newly developed transformation technique to study a number of important questions in plastid biology.

Mark S. Guyer, Ph.D.

References:

1. Klein, T.M., E.D. Wolf, R. Wu, and J.C. Sanford. 1988. High-velocity microprojectiles for delivering nucleic acids into living cells. Nature 327: 70-73.

2. Boynton, J.E., N.W. Gillman, E.H. Harris, J.P. Hosler, A.M. Johnson, A.R. Jones, B.L. Randolph-Anderson, D. Robertson, R.M. Klein, K.B. Shark, and J.C. Sanford. 1988. Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. Science 240: 1534-1538.

3. Johnston, S.A., P.Q. Anziano, K. Shark, J.C. Sanford, and R.A. Butow. 1988. Mitochondrial transformation in yeast by bombardment with microprojectiles. Science: 240: 1538-1541.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 13, 1988

From Chief, Molecular and Medical Genetics Section, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Specific Aims - The Story of Gene Targeting"
R01 GM 21168 and R01 GM 20069 (Capecchi, Mario and Smithies, Oliver)
University of Utah and University of North Carolina

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 12-19-88

For a few NIGMS grantees, the term "specific aims" has taken on a new meaning. These grantees are developing methods for gene targeting in mammalian cells. Gene targeting (also known as targeted integration) allows scientists to introduce foreign genes into chromosomes, not just anywhere, but in specific locations. Although gene targeting is routine in bacteria and yeast, DNA fragments introduced into a mammalian cell integrate into host chromosomes apparently at random. What is needed is a system by which an incoming DNA sequence finds its homologous sequence in the cell's DNA (targeting) and integrates into that site by homologous recombination. The lack of such a system has frustrated geneticists interested in mammalian genetics, especially when they compare the advances in yeast genetics made possible largely by the ease of gene targeting in that organism. Targeted integration of DNA in yeast has proved an enormously powerful tool, allowing scientists to identify DNA sequences that regulate gene expression, to study chromosome structure, and to examine DNA repair, recombination, and replication with a precision that makes mammalian geneticists quite jealous.

Gene targeting has important implications for basic genetic studies, but also can be applied to developing animal models of genetic disorders and to gene therapy. In the past, animal models have been developed by exposing many animals to mutagens and screening for animals that appear to have an analog to a particular human disorder. This approach is extremely unpredictable and laborious and may not produce true analogs to human disorders. Gene targeting is also an essential step toward effective gene therapy, which requires that the correct DNA sequence target and replace the incorrect genomic sequences exactly.

Efforts to target genes in mammalian systems have been thwarted by two problems. First, the frequency of illegitimate (random) integration compared to the desired event is high. Second, the tools for detecting targeted integration have not been available. Recently, two NIGMS supported researchers have devised ingenious schemes both to detect targeted integration and to increase the frequency of homologous (targeted) recombination in cells. This research by Dr. Mario Capecchi and Dr. Oliver Smithies, working independently, seems to promise that gene targeting may soon become a routine tool for mammalian geneticists.

Both Drs. Capecchi and Smithies have taken advantage of improvements in methods for isolating and growing progenitor cells for specific organisms or tissues in culture. Stem cell technology, developed largely by scientists in Europe, is well worked out in mouse embryonic stem cells. Investigators are able to collect embryonic stem cells from mouse blastocysts, culture them on layers of fibroblasts, and reimplant the cells into pseudopregnant females where they develop normally. Stem cell technology is also being explored in rats, sheep, and pigs with the hope that these animals may also prove to be useful biological models. In addition, scientists interested in somatic cell gene therapy may wish to alter cells only in specific tissues, such as bone marrow or liver. The ability to isolate, engineer, and reimplant stem cells from these tissues would allow for modification of specific somatic cells with less risk of undesired effects in other tissues and without modifying the germ line.

Drs. Capecchi and Smithies work with mouse embryonic stem cells, using the technique shown in Figure 1. Stem cells from mouse blastocysts provide the experimental system. Vectors for the incoming DNA are of two types, named for their shapes when drawn schematically. Both vectors contain regions of homology with the target site, and it is at these regions that homologous recombination occurs. Omega vectors, shown on the right, are designed to replace a segment of recipient DNA with incoming DNA. O vectors, shown on the left, insert a DNA segment at a specific site without deleting any of the homologous site. Vectors with sufficient homology to the target sequence are up to 2000 times more likely to locate and integrate into the correct site in the chromosome. Oddly enough, replacement (omega) events and insertion (O) events occur with equal frequency, although replacement requires a double recombination event and insertion requires only a single recombination event.

Once the desired recombination event has occurred, the stem cells are injected into another blastocyst and reimplanted into a pseudopregnant mouse. Chimeric offspring are bred to determine if modified cells have been incorporated into the germ line. If so, the modification is easily maintained by appropriate breeding schemes.

The second problem with targeted integration - detection of a targeted event - has also been addressed by Drs. Capecchi and Smithies; however, their approaches are significantly different. Dr. Capecchi and his colleagues, Kirk Thomas and Suzanne Mansour, have based their approach on a straightforward selection scheme using the bacterial neomycin resistance (neo) gene, which confers resistance to the drug G418. In 1987, they successfully constructed a vector containing both the neo gene from bacteria and an exon from the HPRT gene (in humans, defects in the HPRT gene cause Lesch-Nyhan syndrome). The presence of the HPRT exon targeted the vector to the genomic HPRT gene, where the vector was inserted into the host genome. The host HPRT gene was interrupted by the neo gene and was unable to function properly. Cells in which targeted integration had occurred could be selected in culture on the basis of their survival in the presence of neomycin and their deficiency of HPRT.

Dr. Capecchi has now created new vectors which enable him to study genes which, unlike HPRT, do not have a selectable phenotype. He has succeeded in targeting an oncogene, int-2, which is implicated in developmental regulation. The targeting vector is a cloned version of the int-2 gene and has two markers - the neo gene in the middle and a herpesvirus thymidine kinase (TK) gene at the end. When the vector is integrated into recipient DNA at a random, nontargeted site, the cell becomes resistant to the drug G418 (by virtue of the neo gene) but will be killed by the drug gancyclovir (GANC), which is converted to a toxic by-product by the TK gene. When the vector is integrated into the int-2 gene, however, the TK gene is removed, and the cell becomes GANC resistant. This simple, but elegant, selection scheme allowed Dr. Capecchi to select cells in which the int-2 gene alone has been targeted.

Dr. Smithies has used a different approach, based on the polymerase chain reaction (PCR), to identify cells containing modified DNA. Vectors containing homology to the target DNA are microinjected into mouse embryonic stem cells where they may integrate into the mouse genome by homologous recombination. After cell division, some progeny cells are set aside for possible reimplantation and others are tested for targeted integration by PCR, as shown in Figure 2. Specific sequences in target DNA and the incoming vector are chosen as primers for DNA replication. As the figure illustrates, target DNA and incoming DNA sequences which have not undergone recombination will accumulate at a linear rate; however, molecules in which integration has occurred at the appropriate site will produce a small recombinant sequence at an exponential rate. After twenty or more replication cycles, the presence of this sequence is easily detected by electrophoresis. Once a correctly modified cell has been identified, its sister cells, which were earlier set aside, are collected and reimplanted into a pseudopregnant mouse.

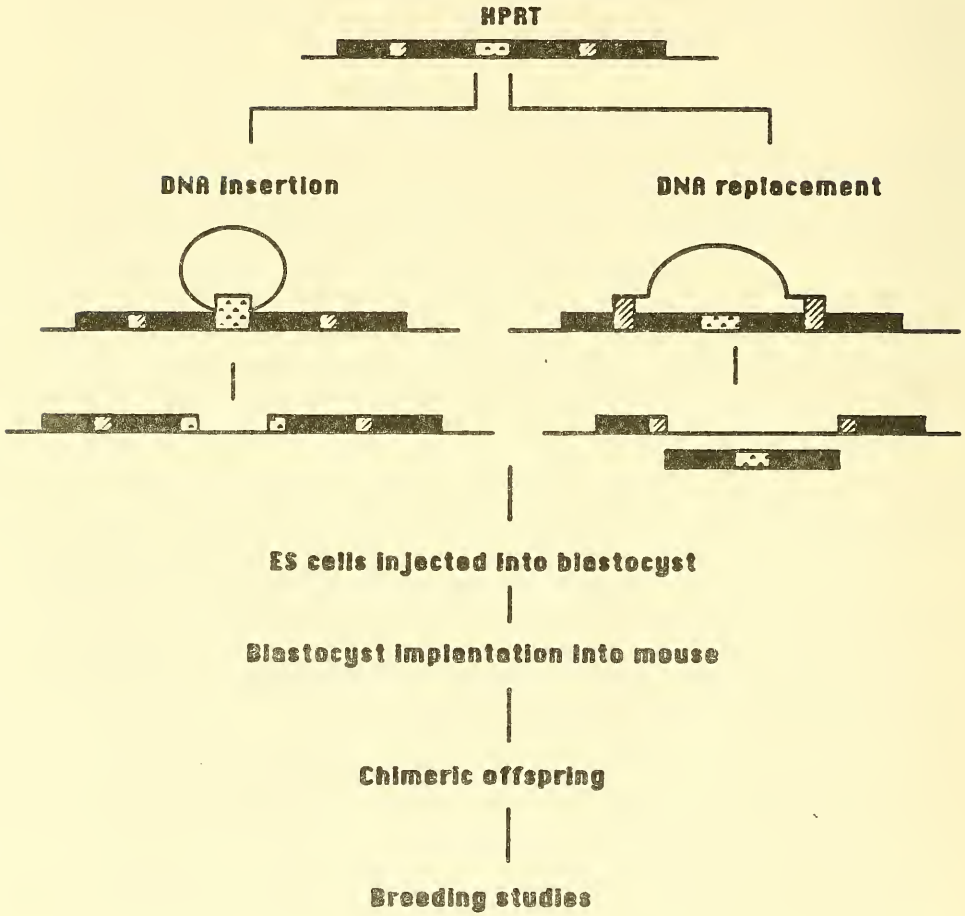
Both investigators have obtained chimeric offspring and will likely soon have experimental mice with genes altered by gene targeting.

The future of gene targeting looks bright. The application of embryonic stem cell technology and the polymerase chain reaction, as well as the design of new vectors, has spurred the field ahead far faster than the expectation of a few years ago. Within the foreseeable future, it should be possible to create animal models, both for studies of human genetic disorders and for basic research on genetic mechanisms in mammalian cells, with precision. Ultimately, research on gene targeting may lead to approaches for gene therapy of human genetic disorders.

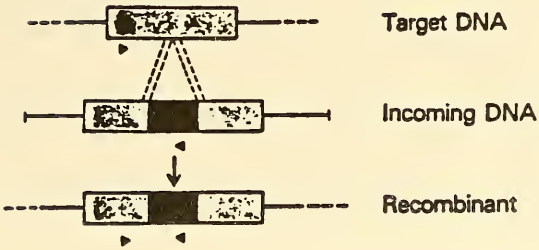
Irene Anne Eckstrand, Ph.D.

GENE TARGETING
USING EMBRYONIC STEM CELLS

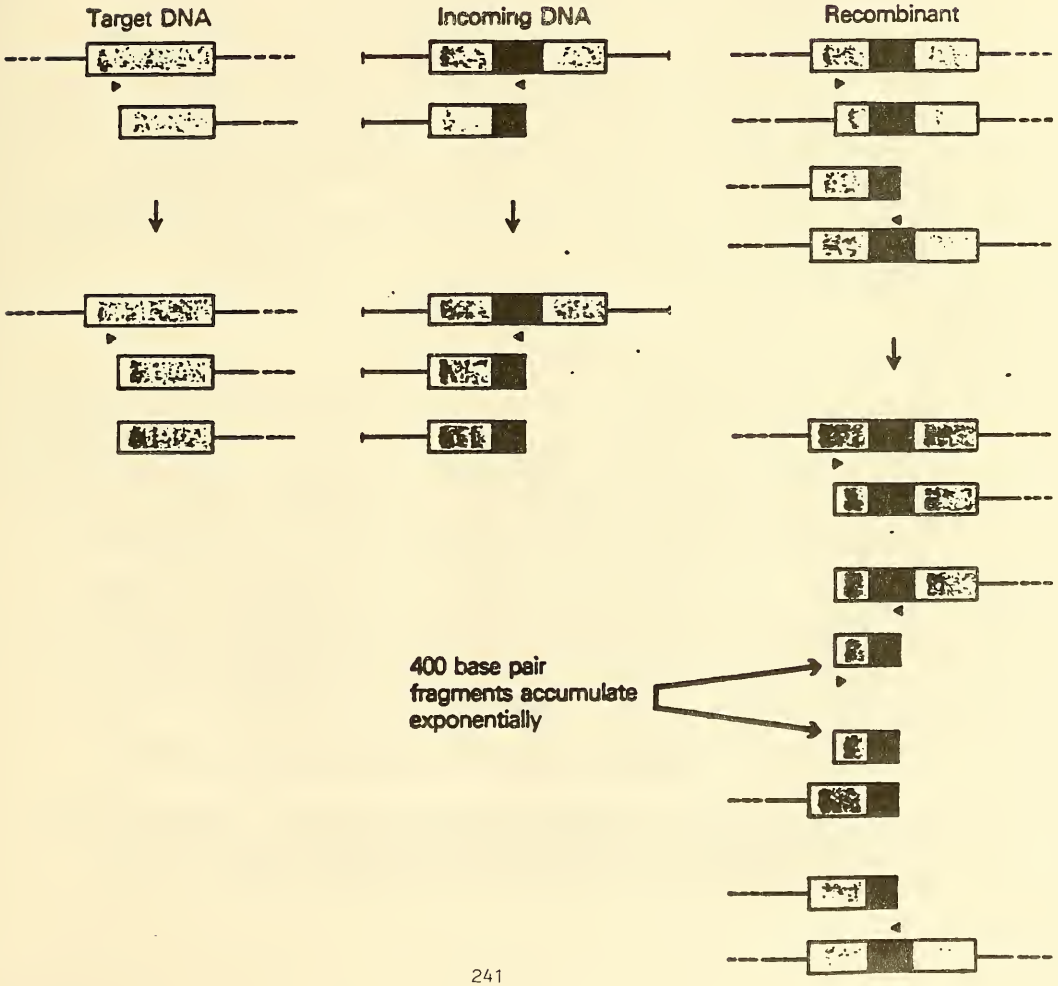
Embryonic stem (ES) cells
derived from mouse blastocysts
provide target DNA



Recombinant Fragment Assay



▶ refers to primer sites



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date June 29, 1988

From Chief, Molecular and Medical Genetics Section, Genetics Program, NIGMS

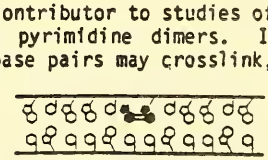
Subject RESEARCH HIGHLIGHT: "Corrections to the Story of DNA Excision Repair"
GMOS901 (Hanawalt, P.), Stanford University

To

Director, NIGMS
Through: Director, Genetics Program

It is usually assumed that the DNA repair process in mammals treats all damaged DNA equally, regardless of its function or location. However, Dr. Philip Hanawalt and his postdoctoral student, Dr. Isabel Mellon, now have evidence that repair of certain lesions in damaged DNA is selective. They have found that transcribed genes are repaired more efficiently than nontranscribed regions of DNA. Further, within a gene, the transcribed DNA strand is repaired more efficiently than is the nontranscribed strand.

Dr. Hanawalt is a major contributor to studies of repair of a specific kind of DNA damage - UV-induced pyrimidine dimers. In the presence of ultraviolet light, adjacent thymine base pairs may crosslink, as shown below.



The presence of these dimers in a DNA helix effectively blocks DNA replication and transcription. Excision repair (one of several DNA repair pathways) is a complex process required for dimer repair. The details of excision repair in mammalian cells have yet to be characterized. In fact, only one human cDNA which encodes a protein with excision repair activity (ERCC-1) has been cloned, and its function is still unknown. However, it is known that many enzymes participate in the excision repair pathway, which is outlined below:

The first step in excision repair involves recognition of damage and incision of the damaged strand near the dimer. Evidence of the complexity of this step comes from studies of xeroderma pigmentosum, a disease caused by defective excision repair. At least nine genetic complementation groups for this disorder are known, suggesting that many enzyme activities are involved in recognition and incision.

Incision is followed by removal of a short stretch of nucleotides, including the damaged DNA.

New DNA is synthesized across the repair patch. The complementary strand presumably serves as a template for the new synthesis. The ends of the repair patch are then ligated to the parental DNA strand.

There is some evidence that excision repair targets certain regions of the genome. For example, mammalian cells that have been UV-irradiated rapidly recover their ability to synthesize RNA, although the overall rate of repair is quite slow. This suggests that active genes are either protected from or recover from DNA damage more rapidly than bulk DNA. Further, data on the rate of repair resulting from UV damage suggest that dimers are removed at two different rates. Dr. Hanawalt thought that excision repair might preferentially repair transcriptionally active genes and that the bulk of DNA might be repaired later or by a different, slower pathway.

To investigate this hypothesis, Dr. Hanawalt chose to work first with Chinese hamster ovary cells which have amplified the gene for dihydrofolate reductase (DHFR) about fifty fold. The objective was to compare DNA repair rates within the DHFR gene to repair rates in neighboring regions. To assay repair, cells were irradiated, then lysed immediately or at various time intervals. DNA was purified and divided into two samples - one untreated and the other treated with T4 endonuclease V, which makes a single strand break at each dimer. The two samples were denatured, electrophoresed in parallel, and exposed to a labeled probe. The amount of hybridization of the probe to the samples was determined by scanning densitometry. When repair efficiency is high, few T4 endonuclease-sensitive sites remain in the DNA, resulting in more full-length restriction fragments. An estimate of the average number of dimers can be calculated from the ratio of full-length restriction fragments in the treated and untreated samples.

Dr. Hanawalt found that 70% of the dimers were removed from a restriction fragment within the gene but that little repair was detectable in a fragment 30 kb upstream. On average, 15% of the dimers were repaired in the total genome in 24 hours. When the experiment was repeated using human cell lines, it was clear that dimers in the DHFR region were repaired earlier than were those in the bulk of the genome. The overall repair proficiency was the same in both regions.

On the basis of these results, Dr. Isabel Mellon, a postdoctoral student in Dr. Hanawalt's laboratory, decided to determine if there was a preference of the excision repair machinery for the transcribed DNA strand. Her results were both unexpected and remarkable. Using probes from the DHFR gene and a technique much the same as that described above, Dr. Mellon found a significant difference in the rate of dimer removal in the transcribed and nontranscribed strands. Within two hours, about 40% of the dimers were removed from the transcribed strand, whereas there was no detectable repair in the nontranscribed strand. By four hours, 70% of the dimers had been removed from the transcribed strand, but less than half were removed from the nontranscribed strand. By eight hours, all of the dimers had been removed from the transcribed strand; only 50% had been removed from the nontranscribed strand. The kinetics of repair in the nontranscribed strand closely resembled the kinetics of repair in bulk DNA.

Dr. Hanawalt's and Dr. Mellon's results offer an explanation for the puzzling observation, mentioned above, that RNA synthesis returns to normal rapidly in irradiated cells, long before significant levels of repair are detected in total DNA. This could lead to an understanding of Cockayne's syndrome, in

which patients are hypersensitive to the lethal effects of UV light. RNA synthesis fails to recover in cells from patients, even though excision repair of dimers in total DNA appears normal. It is possible that the selective repair process, but not the overall repair system, is defective.

Dr. Hanawalt's observations could solve another paradox - that rodent cell lines and human cell lines have similar UV resistance even tnough rodent cells have lower levels of repair. Dr. Hanawalt's results suggest that preferential repair of essential or transcribed genes in the rodent cells might be the explanation.

Strand-specific repair has some important consequences for the cell cycle, not the least of which is the prediction that the number of dimers in any given gene will be different in the two daughters of an irradiated parent cell. Because dimers block DNA replication, mutations should accumulate in daughter cells receiving the unrepaired strand. The statistical expectations of this hypothesis are being examined, so that experiments to study it further can be designed.

Strand-specific repair also has significant implications for mechanisms of excision repair in mammalian cells. Dr. Hanawalt and Dr. Mellon suggest that there may be separate, independent pathways for processing damage in transcribed and nontranscribed sequences and that repair of transcribed strands is directly coupled to transcription. It is possible that premature release of RNA polymerase at the site of a dimer could act as a signal for repair in the transcribed strand. An alternative hypothesis is that the RNA polymerase complex is actively involved in repair.

These studies have many implications for risk assessment. For example, because rodents are widely used in carcinogen testing, it is important that the unique features of DNA repair in rodent and human systems be studied and appreciated. Further, it is now clear that DNA repair in mammals depends on the type of lesion, its precise location in the genome, and the functional state of the DNA at that site. Information about overall DNA damage and repair in one domain of the genome may tell scientists little about what is happening in other regions. Currently, our ways of assessing risk for mutagenic effects may not do justice to the complexities of the repair process.

Irene Anne Eckstrand

Irene Anne Eckstrand, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date June 28, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "A New Connection Between DNA Repair and Ubiquitin"
R01 GM 31530-06 (Varshavsky, A.), Massachusetts Institute of Technology

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 6-28-88

Ubiquitin is a highly conserved cellular protein that serves as a tag to identify other proteins destined for degradation. Although the importance of ubiquitin in cellular metabolism has always been appreciated, the connection between it and other cellular processes was not known until Dr. Alexander Varshavsky began studying ubiquitin several years ago. His studies revealed several surprising and exciting facets of the ubiquitin system. In a study designed to elucidate the role of ubiquitin tagging, Dr. Varshavsky unexpectedly discovered a second protein degradation scheme (see Highlight "Selective Protein Degradation, the N-end Rule," Dec. 30, 1986). Recently (1), Dr. Varshavsky has reported yet another unexpected observation in his continuing study of ubiquitin: the yeast DNA repair gene, rad6, encodes an enzyme that mediates the conjugation of ubiquitin to specific histones.

The yeast rad6 gene is required for a variety of cellular functions including DNA repair, induced mutagenesis and sporulation. Yeast cells containing rad6 mutations are very susceptible to DNA damage induced by agents such as ultraviolet light and display a variety of phenotypes, suggesting that the RAD6 protein is central to a number of regulatory pathways. The connection between the RAD6 protein and ubiquitination of certain histones was found as a result of a systematic study in Dr. Varshavsky's laboratory to identify enzyme activities involved in ubiquitination of proteins.

The enzymatic pathway for protein ubiquitination has been extensively studied with enzymes isolated from mammalian reticulocytes (2). Two enzymatic activities are involved: E1--ubiquitin-activating activity, and E2--ubiquitin-conjugating activity. The initial phase in Dr. Varshavsky's study of ubiquitination was to isolate these two activities from yeast by chromatography of cell extracts on a ubiquitin-Sepharose column. Both E1 and E2 enzyme activities bound to the column and were selectively eluted using a reducing agent, dithiothreitol (DTT). Gel filtration of the DTT eluate resulted in the separation of the E1 and E2 activities. One E2 species was of particular interest, because it specifically ubiquitinated histones H2A and H2B in vitro in the presence of ATP and purified E1. The activity was traced to a 20,000-dalton protein.

Dr. Varshavsky's group proceeded with the cloning of the genes for both activities. The gene for the 20,000-dalton ubiquitin-conjugating (E2) protein was isolated by determination of the partial amino acid sequence of the purified protein and by screening a genomic library with a synthetic DNA probe corresponding to the coding sequence of the known protein. The 14-residue amino acid stretch that was determined during this process was used to search a protein sequence database for known proteins with sequence similarities. This revealed that the protein encoded by the rad6 gene contained an amino acid sequence identical to that of the 20,000-dalton E2 protein. In addition, the known molecular weight for the RAD6 protein was 19,700 daltons, very close to the apparent molecular weight of E2. The identity of the isolated E2 gene was verified by restriction mapping and comparison to the known restriction map for the yeast rad6 gene.

In order to demonstrate that the rad6 gene was indeed the E2 (ubiquitin-conjugating) activity in the yeast system, a 600 base pair segment of DNA containing the cloned rad6 gene was inserted into an expression vector and transformed into E. coli cells. Extracts from cultures expressing the rad6 gene and from control cultures were assayed for the presence of ubiquitin-conjugating activity as measured by the conjugation of ^{125}I-labelled ubiquitin to histone H2B. A high level of H2B-ubiquitination was found in the extracts from cells containing the rad6 gene, demonstrating that the RAD6 protein was identical to the E2 activity in yeast. These results confirmed that the RAD6 protein is encoded by a member of a group of closely related genes, which Dr. Varshavsky calls the UBC (ubiquitin-conjugating enzymes) genes. Another example of an enzyme from this family has come from recent work in Dr. Breck Byers' laboratory (GM 18541). A yeast cell division cycle gene, CDC34, has been identified that encodes a protein with sequence homology to the RAD6 protein, as well as to cloned E2 proteins from other species.

The next question to ask was, what role does the ubiquitination of specific histones play in DNA repair? Earlier in vitro experiments in several laboratories demonstrated the importance of chromatin in regulating the accessibility of damaged DNA to repair enzymes. Thus, two possible roles of ubiquitination of histones in repair can be imagined. First, the ubiquitin tagging of histones (and possibly other basic proteins) in a damaged region of the DNA might lead to the degradation of the proteins and the exposure of the damaged DNA to repair enzymes. Alternatively, ubiquitination of chromosomal proteins could directly induce structural changes in the chromatin, thereby permitting access of repair enzymes to the damaged DNA. Targeting of the RAD6 protein to the site in need of DNA repair could be related to selective exposure of that site as a result of DNA damage. Transient ubiquitination of chromosomal proteins might also be required for other events requiring exogenous enzyme access to DNA such as DNA transcription, recombination and replication. Such a model is supported since there is a disproportionate amount of ubiquitinated chromatin in stretches of drosophila and mammalian DNA that have the potential to be highly expressed.

Given the apparent involvement of ubiquitin in metabolic processes and its conservation across species, it is possible that known metabolic diseases caused by defects in repair activity, such as xeroderma pigmentosum, may be

due to defects in the ubiquitin-conjugating enzymes. Dr. Varshavsky plans to examine a number of human cell lines that are known to have repair defects, to determine if changes in the E2 enzyme have caused the defects. Clearly, the ubiquitin system is diverse and will most likely yield more surprises as Dr. Varshavsky continues his studies.

Jane L. Peterson, Ph.D.

References:

1. Jentsch, S., McGrath, J., and Varshavsky, A. (1987). The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme. Nature 329, 131-134.

2. Hershko, A. and Ciechanover, A. (1986). Prog. Nucleic Acid Res. Molec. Biol.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date February 15, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Telomere Replication"
R01 GM 26259-11 (Blackburn, E.), University of California, Berkeley
R01 GM 19199-17 (Prescott, D.), University of Colorado

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 2/23-89

The ends of chromosomes, called telomeres, are unique structures whose primary role is to provide stability for the chromosome. The replication of telomeres has been of interest because usual models for DNA replication do not account for how the linear ends of chromosomes are duplicated. One problem in studying telomere duplication has been the difficulty in obtaining large numbers of identical chromosome ends. However, several unicellular organisms, such as Tetrahymena and Oxytricha, have thousands of identical small chromosomes in addition to a micronucleus that contains the cell's complete genetic information. These small chromosomes make up the macronucleus that is present during the growth phase of the life cycle. Two NIGMS grantees, Dr. Elizabeth Blackburn and Dr. David Prescott, have been studying telomeres in Tetrahymena and Oxytricha. They have examined telomere structure and replication in both organisms and found that they share a similar overall architecture that differs only in the details.

In Tetrahymena, rDNA is amplified thousands of times and maintained in the macronucleus as linear minichromosomes, an excellent source of identical telomeres for study. Dr. Blackburn's studies in Tetrahymena have demonstrated that: a) telomeres are dynamic rather than static structures, as once thought (see Highlight, "Telomeres: Never-ending Structures?", July 23, 1986); b) the DNA in the telomeres contains a random number of sequential repeats (5' TTGGGG 3'); c) the number of DNA repeats varies with the life cycle of the organism; d) single-stranded DNA breaks of unknown function have been detected in the repeats at the end of the telomere; and e) telomere elongation occurs by addition of nucleotides to the end of the chromosome in the absence of a DNA template. In her continuing investigation of this system, Dr. Blackburn has now further characterized telomere elongation and showed that an unusual ribonucleoprotein, which she has named telomerase, is the enzyme responsible for adding terminal nucleotides to telomeres (1).

To define the DNA sequence needed to prime elongation at the telomere end, Dr. Blackburn added DNA fragments, whose base compositions were different, into crude Tetrahymena extracts which had telomerase activity. She found that telomere repeats (5'TTGGG3') were added only to G-rich fragments. If the fragment did not contain four terminal G's, additional G's (up to a total of four) were added, before the repeat sequence was added. All the sequences that served as primers shared the ability to form self-associating

G-G base-paired hairpin structures under nondenaturing conditions (2). These results suggested that telomerase activity is not dependent upon recognition of a specific sequence but rather upon the presence of any G-rich sequence capable of forming a hairpin structure. The requirement that an enzyme recognize a secondary structure is not unusual, but the addition of nucleotides to a terminus to form the proper structure is novel.

The next step in studying telomerase was to purify the enzyme. Gel filtration of a crude extract of Tetrahymena yielded an active fraction with a molecular weight of 200-500 kilodaltons. Treatment of the fraction with micrococcal nuclease, RNase or proteinase destroyed the activity, indicating that the enzyme was an RNA/protein complex. After further extensive purification by gel filtration, two small RNA species, 154 and 80 nucleotides in size, were isolated. These RNAs contain the sequence CAACCCCAA, which, because of its complementarity, could serve as the template for synthesis of the TTGGGG terminal repeat (3).

Dr. Blackburn's laboratory has proposed a model in which telomerase recognizes the end of the telomere DNA by the complementarity between its RNA and the single-stranded telomere tail. The RNA next serves as a template for the addition of several more nucleotides to the telomere DNA. Further elongation occurs by translocation of the enzyme to the 3' end of the DNA (figure 1). If this model is correct, telomerase would be unique among eukaryote ribonuclear proteins in catalyzing DNA synthesis using an RNA template.

Following the protocols developed in Dr. Blackburn's laboratory, Dr. Prescott has detected a telomerase-like activity in Oxytricha (4). Oxytricha's macronuclear genome contains 10,000 gene-sized pieces of DNA that range in size from 400 to 20,000 base pairs. Oxytricha telomerase activity in vitro adds from two to seven copies of GGGGTTTT (the Oxytricha terminal repeat) to the end of the genes. Unlike Tetrahymena, the telomere length in Oxytricha is invariably 36 nucleotides in length. As addition of the terminal repeats proceeds, the reaction appears to pause after the fourth T, suggesting that the telomerase may have to dissociate and reassociate on the oligonucleotide after each addition. Such pausing would be consistent with translocation of the enzyme during elongation, as predicted in Dr. Blackburn's model.

The uniform size of the Oxytricha telomere may be explained if the ends are synthesized in oversized lengths and then trimmed to a uniform size. Dr. Prescott proposes that proteins bind to the terminus at a specific site (at nucleotide #36) in the Oxytricha telomere. Then, a nuclease, such as an uncharacterized activity he has consistently detected in Oxytricha extracts, acts to clip the excess single-stranded DNA, resulting in uniformly sized telomeres. In fact, Dr. Daniel Gottschling, a postdoctoral fellow (F32 GM 10442) in Dr. Virginia Zakian's laboratory, has successfully isolated two proteins that bind specifically to the 3' extension of Oxytricha telomeres (5) and Dr. Thomas Cech's laboratory (GM 28039) has shown that these proteins give the Oxytricha terminus a distinctive pattern of protection from methylation (6).

Dr. Prescott has proposed a replication model for the Oxytricha minichromosomes (Figure 2), in which the single-stranded DNA terminus serves as a template for an RNA primer which in turn serves as a primer for DNA replication. Although the model is not based on experimental evidence, all of the enzyme activities involved in the model have been reported except for the RNA primase. An unusual feature of the model is that two replication forks would have to pass each other at some point during replication. No model for DNA replication described to date has involved such a mechanism, nor is it apparent how such a mechanism could occur. However, continued studies on the replication of chromosome ends in these simplified systems will give us valuable information about the replication of the ends of linear DNA molecules and will help to explain the function of telomeres in all eukaryotic cells.

Jane L. Peterson

Jane L. Peterson, Ph.D.

References

1. Greider, W. W. and Blackburn, E.H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell 51, 887-898.

2. Henderson, E., Hardin, C.C., Walk, S.K., Tinoco, I., Jr., and Blackburn, E.H. (1987). Telomeric DNA oligonucleotides from novel intramolecular structures containing guanine-guanine base pairs. Cell 51, 898-908.

3. Greider, C.W. and Blackburn, E.H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature 337, 331-337.

4. Zahler, A.M. and Prescott, D.M. (1988). Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear molecules. Nucleic Acid Res., in press.

5. Gottschling, D.E. and Zakian, V.A. (1988). Advances in Cell Biology, Vol. 2, K.R. Miller, ed. (JAI Press), in press.

6. Price, C.M. and Cech, T.R. (1987). Genes and Development 1, 783-793.

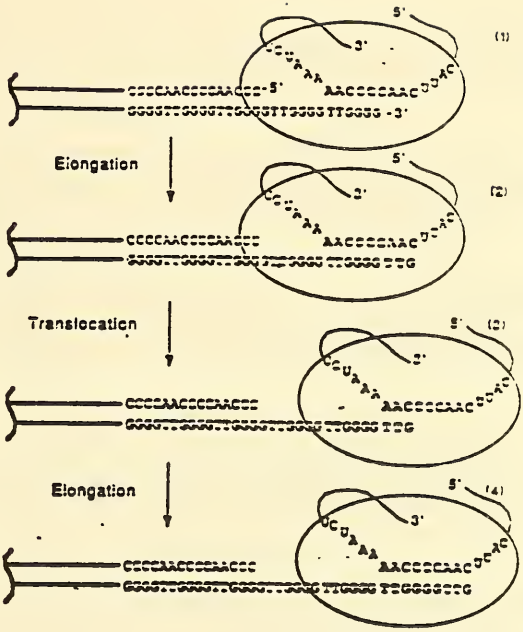


Fig. 1. Model for elongation of telomeres by telomerase. The *Tetrahymena* telomere is shown containing a 13-base overhanging TTGGGG strand[14]. (1) After recognition of the TTGGGG strand by telomerase, the 3' most nucleotides are hybridized to the CAACCCCAA sequence in the RNA. (2) The sequence TTG is then added one nucleotide at a time[4]. (3) Translocation then repositions the 3' end of the TTGGGG strand such that the 3' most TTG nucleotides are hybridized to the RNA component of telomerase. (4) Elongation occurs again, copying the template sequence to complete the TTGGGGTTG sequence. This mechanism explains how oligonucleotides with 3' ends terminating at any nucleotide within the sequence TTGGGG are correctly elongated to yield perfect tandem repeats of $(TTGGGG)_n$.

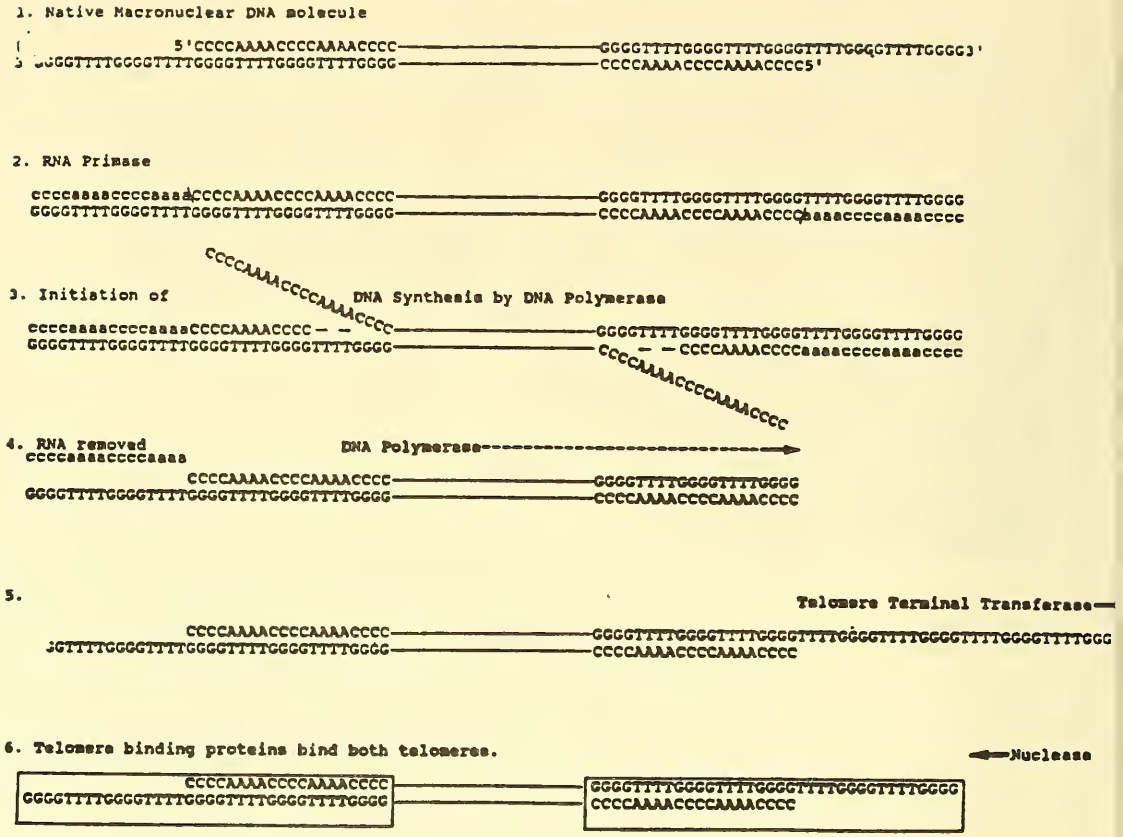


Figure 2 Model for Telomere Replication in the Macronucleus of Hypotrichous Ciliates

1. Native macronuclear DNA molecule with telomere binding proteins removed prior to DNA replication.
2. RNA primer (lower case letters) is made on 3' terminal extension (5' gap) by an RNA primase activity.
3. DNA synthesis by DNA polymerase is initiated by RNA primer.
4. DNA polymerase replicates entire macronuclear gene but leaves a blunt end on the right end of the molecule because there is no template to create a new 3' terminal extension. RNA primer is removed.
5. Telomere terminal transferase makes a new 3' terminal extension whose length is overly long.
6. Telomere binding proteins bind to telomere with overly long single-stranded 3' terminal extension and protects first 16 bases of single-stranded region. Nuclease activity removes unprotected single-stranded region. Replication is complete.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 20, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Linking Maps and Hypotheses"
R01 GM 40886 (White, R.), University of Utah

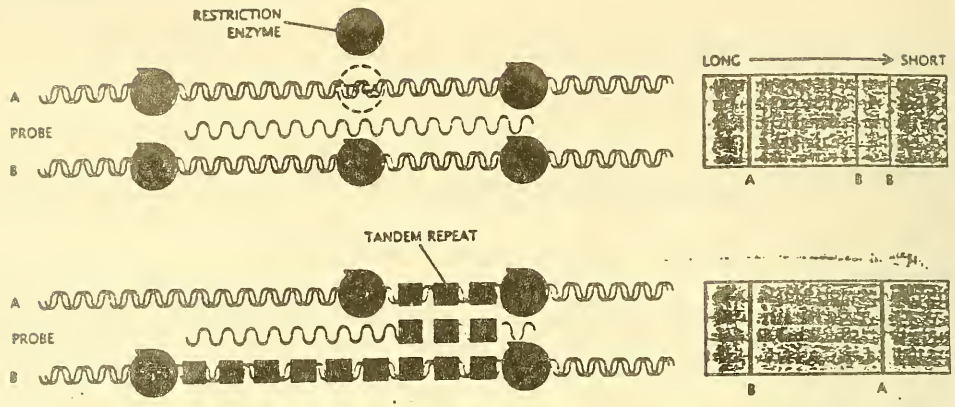
To Director, NIGMS
Through: Director, Genetics Program, NIGMS 3-20-89

Mapping the human genome is a daunting, but essential, first step in the molecular characterization of the human. Finding the 50,000 to 100,000 genes estimated to be present in the human genome and identifying their locations requires a good genetic map. However, it escapes nobody's attention that the map in itself reveals relatively little about the genetic and molecular processes we ultimately hope to understand. Fortunately, we will not have to wait until the map is complete before we can use it to address fundamental biological questions. Scientists are already finding, within the rudimentary map now available, intriguing and interesting clues to how the human genome works.

Dr. Ray White is a major contributor to the current human linkage map. He has also been alert to features of the map that can be used to address biological questions. For example, recent information from his laboratory suggests that a long-held notion about how variation in the human genome and other genomes arises may be incorrect. Specifically, Dr. White's data suggest that variation in the numbers of tandemly repeated DNA sequences may not be generated by unequal crossing over of chromosomes during meiosis.

As part of his program to develop a linkage map of the human genome, Dr. White searches for DNA markers that may be associated with specific genes and will therefore be useful for placing these genes on the map. The value of any DNA marker depends largely on how many variants are present in the population. The more variants there are, the more likely it is that an individual with a specific disease gene will have two different forms of the marker, making it feasible to detect recombination between the disease gene and the marker. In 1980, Dr. Arlene Wyman, working in Dr. White's laboratory, discovered that certain regions in the genome are hypervariable; some have 70 or more variants which can be detected by the sizes of restriction fragments. It turned out that the hypervariable regions are constructed of tandemly repeated DNA sequences. The number of copies of the repeated sequence varies widely among individuals. These variable number tandemly repeated sequences, called VNTRs, are easily detected by any restriction enzyme whose recognition sequence is not present in the VNTR.

The VNTR system is compared to the more common RFLP marker system below:



DNA MARKERS—sites at which homologous chromosomes often differ in DNA sequence—are detected as RFLP's (restriction-fragment length polymorphisms). The DNA is digested with a restriction enzyme, which cuts wherever it finds a specific short sequence of nucleotides (in this case the base sequence *TCGA*). In one kind of marker (*top left*) a sequence difference causes a restriction site that is present on one homologous chromosome to be absent from the other. As a result the restriction fragments from each chromosome will differ in length. A DNA probe whose base sequence is complementary to that of DNA at the marker site reveals the fragments after they are sorted by electrophoresis (*top right*). Another kind of marker (*bottom left*) is characterized by a VNTR—a variation in the number of tandem repeats (short, repeated DNA sequences). The span between cutting sites differs between matching chromosomes, again resulting in distinctive fragments detected after electrophoresis (*bottom right*).

In 1985, Jeffreys reported that VNTR regions in dispersed regions of the genome had related sequences. This finding enabled other researchers to build synthetic oligonucleotide probes which could be used to detect many of the VNTRs simultaneously. Jeffreys' probes have become valuable resources for genetic linkage mapping, for forensic identification, for paternity testing, and for determining the likelihood of success of tissue transplants.

The existence of VNTRs, while fortuitous for the above reasons, raises fundamental questions about the organization and behavior of the genome. For example, how are the basic units added and subtracted from VNTR regions? The most widely accepted hypothesis is that unequal crossing over during meiosis generates tandem repeats. This hypothesis has been especially attractive because the repeat sequences bear a fairly strong similarity to chi, which is a recombinational hotspot in E. coli (Jeffreys et al, 1985), and to a hotspot in the mouse genome (Steinmetz et al, 1986).

Dr. White noted that a prediction of the unequal crossing over hypothesis is that flanking markers surrounding a VNTR will be recombinant following an event that generates a new allele. By searching families in the many three-generation pedigrees collected for linkage analysis, Dr. White found a family in which a new allele had arisen. In this case, the father's allele

had been inherited intact, but a new allele, smaller than either of the mother's alleles, appeared in a daughter. It was natural, then, to inquire whether or not the new allele had arisen by unequal crossing over in a maternal meiotic division.

Cosmid libraries of the DNA from both mother and daughter were constructed so that DNA sequences on each side of the VNTRs could be determined. 750,000 clones were screened to obtain the VNTR and flanking sequences for both of the maternal sequences and the single altered daughter sequence. The DNA sequences of these fragments, approximately 650 base pairs apiece, were then determined. The sequence data clearly showed which maternal chromosome had been inherited by the daughter; one of the mother's chromosomes was identical to the daughter's chromosome except for the loss of exactly one repeat unit. Surprisingly, the DNA sequences on each side of the VNTR regions on those two chromosomes were also identical. If the unequal crossing over hypothesis were correct, one would expect that the sequence on one side of the VNTR on the daughter's chromosome would be identical to one maternal chromosome while the sequence on the other side would correspond to the sequence of the other maternal chromosome.

These data do not exclude the possibilities of two independent crossing over events within the very short region between the VNTR and the flanking sequences or of an unusual gene conversion event. However, such explanations fail to account for other features of the VNTR region. It is more likely that some mechanism other than unequal crossing over is responsible for generating variation in the numbers of tandemly repeated sequences.

What that mechanism might be is still undetermined. What is interesting and encouraging is that work focused on fairly straightforward genetic linkage mapping has resulted in some very important hypothesis testing. It is clear that developing a resource of genomic information need not take place in the absence of excellent science. In fact, one would hope that they would occur in tandem.

Irene A. Eckstrand, Ph.D.

References

Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985). Hypervariable "minisatellite" regions in human DNA. Nature 314:67-73.

Steinmetz, M., Stephan, D., and Lindahl, K. F. (1986). Gene organization and recombinational hotspots in the murine major histocompatiblity complex. Cell 44:895-904.

White, R. and Lalouel, J. M. (1988). Chromosome mapping with DNA markers. Scientific American 258:40-48.

Wolff, R. K., Nakamura, K., and White, R. (1988). Molecular characterization of a spontaneously generated new allele at a VNTR locus: No exchange of flanking DNA sequence. Genomics 3:347-351.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date April 4, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Inversions and Recombinations" T32 GM 07464, (Sherwood Casjens, Program Director) and R01 GM 27068, (Roth, John) University of Utah

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 4/4-4-89

Escherichia coli and Salmonella typhimurium are enteric bacteria frequently used as models for studies on a variety of biologic phenomena. The two genera are believed to have diverged over 150 million years ago. Their DNA sequences are significantly different and little, if any, exchange of DNA has occurred between the species. Despite major differences between the organisms, the order of genes on the two bacterial chromosomes is remarkably similar. Are there evolutionary advantages to this gene sequence or are the conserved sequences the result of mechanical inhibitions against rearrangement? These questions were explored by NIGMS grantee John Roth and two of his graduate students, Anca Segall and Michael Mahan, both supported by the Genetics Predoctoral Training Grant at the University of Utah.

To address this issue, Dr. Roth and his students asked the following question: If a chromosome is engineered such that recombination occurs with high frequency between two distant sites, what happens to the order of the intervening genes? Does it remain the same or does it invert? Reciprocal recombination with looping-out (figure 1) will invert the order of the intervening genes whereas non-reciprocal or double recombination (e.g., figure 2) will not.

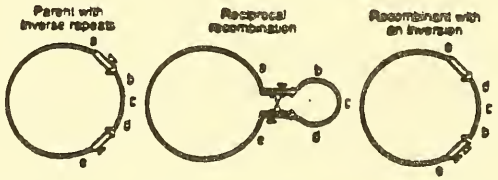


Figure 1. Inversions formed by intrachromosomal recombination. Sequences positioned in the circular bacterial chromosome in inverse orientation can recombine to invert the intervening chromosomal segment. Two copies of the gene (open arrows) each contain a different mutation (triangles). Reciprocal recombination produces a functional gene, which can be selected, and inverts the order of the intervening genes.

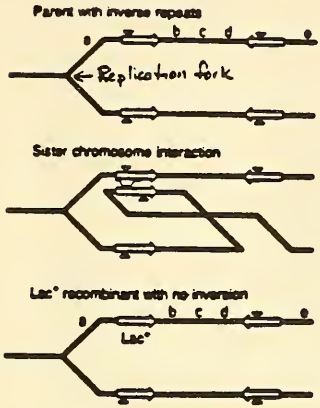


Figure 2. One example of recombination without inversion. Inversely oriented sequences present on sister chromosomes can interact non-reciprocally or by two reciprocal exchanges to generate a functional gene without inverting the order of the intervening genes.

Dr. Roth and his students developed a clever selection scheme to identify rearrangements in salmonella chromosomes. They engineered a bacterial chromosome that contained two copies of a test gene separated by a piece of "normal" bacterial chromosome. Each copy of the gene contained a different mutation, and the normal gene product could be produced only if recombination occurred between the two mutant genes, creating one functional copy and one with both mutations (see figure 1 & 2). In order to enhance recombination between the two copies of the gene, direct repeats of a transposon were attached to one end of each gene. These identical sequences could line up, promoting the crossover that would produce a functional gene. The researchers could identify bacteria containing the functional test gene by selective growth conditions and then determine the gene order between the two gene copies in the recombinants. The scientists used two systems, one involving the utilization of lactose (*lacZ* gene) and one involving histidine synthesis (*his* gene), thus insuring that the findings were not unique to one gene.

The researchers constructed bacterial chromosomes with the two gene copies at many different sites throughout the chromosome and examined thousands of recombinants. The chromosomal intervals between the two mutant genes fell into two distinct classes: permissive intervals had inversions 25 to 90 percent of the time whereas non-permissive intervals had none. Importantly, recombination frequencies were the same in both permissive and non-permissive regions. The two types of intervals were not randomly distributed (see figure 3). Neither the sequences involved in recombination nor those immediately adjacent to them determined whether the intervening DNA would invert. Finally, non-permissive intervals were found within permissive regions.

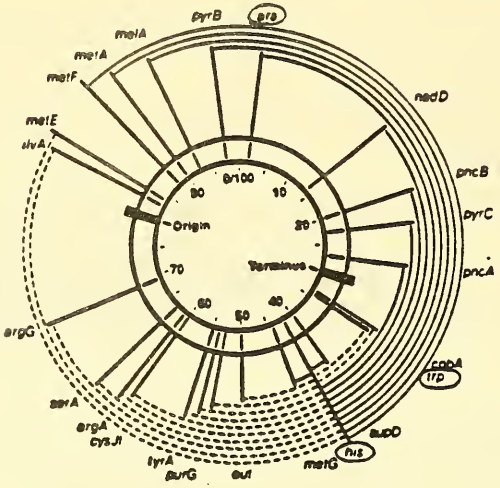
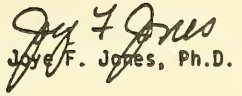


Figure 3. Recombination between his sequences. Solid lines indicate permissive intervals where chromosomal segments invert during recombination. Dotted lines are the non-permissive intervals where inversions have never been found. (The map units, shown around the innermost circle, are minutes, reflecting the time it takes that part of the chromosome to be transferred to a recipient bacterium during conjugation.) Similar results were obtained with lacZ sequences.

Although recombination occurs with equal frequency throughout the chromosome, there are clearly regions where inversions rarely occur. This, however, does not address the mechanism. An inversion could be lethal, and consequently undetectable. If this were true, then one would expect to find numerous recombinants with growth deficiencies due to sublethal replication defects. Of the thousands of recombinants they studied, the researchers found only one (between ara at 2 min on the map and trp at 34 min) that grew more slowly than normal. Thus, it is unlikely that the constraints on inversions are, to any significant extent, due to detrimental effects of those chromosomal rearrangements. Could the lack of inversion be due to physical constraints? The trp (34 min)-his (42 min) interval is always non-permissive (over 2000 recombinants were studied). The scientists engineered an artificial inversion by removing this intervening DNA and reinserting it in the opposite orientation. These cells grew well and the interval did not revert to "normal" sequence. Thus, Dr. Roth and his student colleagues believe that local mechanical constraints are the basis for the non-permissive regions.

The bacterial chromosome may have two functional regions. The entire chromosome may be capable of exchange between sister chromosomes during replication, resulting in recombination without inversion figure 2). If such recombination were slightly aberrant, gene duplication could result. Indeed, gene duplication occurs with equal frequency in permissive and non-permissive regions. In the non-permissive regions of the chromosome, the looping-out necessary for recombination with inversion (figure 1) is inhibited, perhaps because of differential supercoiling or because of a rigid three-dimisional structure in this area. An intriguing observation not addressed by the hypothesis is the non-randomness of the two types of intervals. The permissive region lies on one side of the circular chromosome, the non-permissive region on the other.

Among higher organisms, the order of genes along homologous chromosomes is not necessarily conserved. However, there are regions containing clusters of unrelated genes whose order is very similar among different species. These genes are presumed to have moved as a group during the inevitable chromosomal duplications, inversions, and translocations that occurred during evolution, and it is surprising that the genes did not separate. Dr. Roth's work suggests that mechanical inhibitions may have played a role in keeping these genes together.



Joe F. Jones, Ph.D.

Reference

Segall, A., M.J. Mahan, and J.R. Roth. 1988. Rearrangement of the bacterial chromosome: Forbidden inversions. Science, 241: 1314-1318.

B. MECHANICS OF GENE EXPRESSION AND CONTROL

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date April 4, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "To Change your Sigma, Change your Genes" R01 GM 18568, (Losick, Richard), Harvard University

To Director, NIGMS
Through: Director, Genetics Program, 4-4-89

Development involves an exquisitely orchestrated series of events in which genes are turned on and off in a carefully regulated pattern. Because development is so complex, some very "simple" organisms are being used to study how gene expression is affected by developmental cues. One such system is the development of spores by the gram-positive bacterium, Bacillus subtilis. Bacilli, and other bacteria, live and divide under "normal" conditions of temperature, humidity, and nutrition. When the going gets tough, some genera of bacteria develop spores which are unaffected by the harsh environments that would kill vegetative cells. When normal conditions return, the spores develop into the common rod-shaped bacteria.

During sporulation in B. subtilis, the organism is partitioned into two unequal compartments: the mother cell, which will subsequently die, and the forespore, which will become the mature spore and--like germline cells in metazoans--will give rise to progeny. Different genes are expressed in the two compartments of the sporulating bacillus. This differential gene expression is due to differential transcription which in turn is due to different DNA-dependent RNA polymerases, the enzymes involved in the production of RNA from the DNA template.

Bacterial RNA polymerases consist of a core enzyme (containing three protein subunits) and a sigma factor. The core enzyme catalyzes transcription whereas the sigma factor enhances polymerase binding at the gene promoter. Once transcription begins, the sigma factor dissociates from the enzyme. Different sigma factors bind to different promoter sequences, and B. subtilis uses this system to alter the sets of genes transcribed during different stages of its life cycle. In the forespore, the polymerase uses sigma factor G, whereas the mother cell uses sigma factor K, recently described by NIGMS grantee Richard Losick, his student, Barbara Kunkel, and his postdoctoral fellow, Lee Kroos.

Researchers in Dr. Losick's laboratory were interested in transcriptional regulation of two spore-related genes, spoIVCB, which is activated early in the development of the mother cell, and cotD, a late gene encoding a spore coat protein. They identified a single sigma factor of 27 kilodaltons (kD) which initiated transcription in both genes although cotD was transcribed more strongly than spoIVCB. Closer analysis of the protein mixture in which the new sigma K had been identified revealed a 14kD switch protein that reversed this transcription "preference" by sigma K; with sigma K plus the switch protein, spoIVCB was transcribed more strongly than cotD.

A candidate gene for either sigma K or the switch protein was spoIIID, which Dr. Losick's group had shown to be a direct regulator of spoIVCB. The switch protein is most likely encoded by spoIIID, since its DNA sequence predicts a protein gene product of 93 amino acids, and the first 34 amino acids exactly correspond to the switch protein sequence. Surprisingly, the predicted product of the spoIVCB gene corresponds to the starting sequence of sigma K. In the predicted coding sequence, the amino end of the sigma factor is preceded by 20 amino acids, suggesting that the primary gene product is a "pro-protein" that is processed to the mature sigma factor by removing the initial 20 amino acids. However, the gene's entire open reading frame can encode a protein (including the leader sequence) of only 17.3kD and a mature protein of 15.2kD (135 amino acids); sigma K has 155 amino acids. Where are the last 20 amino acids?

The candidate was another early sporulation gene, spoIIIC, because its gene product had a predicted sequence very similar to that of the carboxy end of other known sigma factors. Using DNA probes specific for spoIVCB and spoIIIC, Dr. Losick and colleagues showed that the two genes are contiguous in mother cells, but are more than 10 kilobases apart in vegetative cells. They are joined not at their ends but within the spoIVCB gene, approximately 112 codons (336 bases) from the beginning of the coding region. The chromosomal rearrangement is first detectable three hours after the beginning of sporulation. This rearranged gene, called sigK, is initially transcribed at a low level by some other sigma factor. Once activated, sigK is self-regulated, since RNA polymerase with sigma factor K transcribes the sigK gene in the presence of the switch protein. The chromosomal rearrangement is apparently irreversible, but since the mother cell is terminally differentiated and destined to die, the rearrangement has no effect on the next generation of bacteria.

Chromosomal rearrangement has been described in a number of systems, most notably in mammalian lymphocytes where it give rise to the nearly infinite array of antibody molecules an animal can produce. Another important example is the developmentally related rearrangement of the nitrogen-fixation genes in cyanobacteria. The rearrangement in B. subtilis is different from these, however, because the rearranged gene is a regulatory gene that directs gene expression within the cell-type in which the rearrangement occurs. This opens up a new way of considering how differential gene expression can occur.

Joy F. Jones, Ph.D.

References:

1. Stragier, P., B. Kunkel, L. Kroos, and R. Losick. 1989. Chromosomal rearrangement generating a composite gene for a developmental transcription factor. Science 243, 507-512.

2. Kroos, L,. B. Kunkel, and R. Losick. 1989. Switch protein alters specifity of RNA polymerase containing a compartment-specific sigma factor. Science 243, 526-529.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date July 24, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Turn Off with Oncogenes" R01 GM 26444-11 (Evans, R.), Salk Institute, La Jolla, CA

To Director, NIGMS
Through: Director, Genetics Program 7-24-89

Oncogenes from tumorigenic viruses can cause neoplastic transformation in virus-infected cells. Similar cellular genes, called proto-oncogenes, are found in normal cells and are not tumorigenic; their gene products function as regulators of cell growth and differentiation. One viral oncogene, v-erbA, which acts synergistically with several other oncogenes to cause transformation in avian erythrocytes, and the related proto-oncogene, c-erbA, have been studied extensively by NIGMS grantee Ronald Evans of the Salk Institute. Dr. Evans and his colleagues showed that c-erbA encodes the thyroid hormone receptor. This receptor is one of a growing family of ligand-modulated, trans-acting gene-regulatory factors which includes receptors for steroid hormones, retinoic acid, and vitamin D. Each receptor has a DNA-binding domain and a ligand-binding domain that function independently. The ligand (hormone or other substance) enters the cytoplasm and binds to the receptor; the complex then binds DNA at sites specific for each receptor, altering the rate of transcription of ligand-regulated genes.

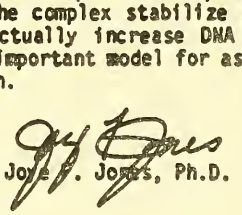
Although c-erbA and v-erbA have very similar DNA sequences, their protein products differ functionally; both bind DNA but only the c-erbA product (the thyroid hormone receptor) binds thyroid hormone. Because oncogenes and proto-oncogenes are so similar, scientists have suggested that oncogenes cause neoplastic transformation by acting as constitutive "on" switches for genes normally regulated by proto-oncogenes. In recently reported work, Dr. Evans challenged this notion. He showed that the thyroid hormone receptor and v-erbA can each act as a negative gene regulator and suggested that the oncogene may act as a constitutive "off" switch.

Genes regulated by thyroid hormone via the receptor-hormone complex contain a thyroid-responsive element, an enhancer-like palindromic DNA sequence to which the receptor binds. To study hormonal regulation, Dr. Evans and his colleagues attached a thyroid-responsive element to a reporter gene encoding an easily measured protein, chloramphenicol acetyltransferase (CAT), thereby bringing the reporter gene under thyroid-hormone receptor control. The engineered reporter gene was inserted into cells along with a test gene, either the gene for thyroid hormone receptor (the c-erbA proto-oncogene) or the v-erbA oncogene. The recipient cell transcribed the test gene and translated the mRNA into the protein gene product; this protein, in turn, regulated expression of the CAT reporter gene via the thyroid-responsive element. All cells receiving the CAT gene would produce CAT. If the test gene enhanced transcription, cells would produce more CAT than controls without the test gene, and, conversely, if the test gene inhibited transcription, cells would produce less CAT than controls. (The transcription level of an unrelated gene was used as an internal control to correct for inherent differences among cells.)

As expected, cells containing the reporter gene plus the c-erbA proto-oncogene produced large amounts of CAT when thyroid hormone was added to the cells. Surprisingly, however, in the absence of the hormone, these cells produced much less CAT than controls lacking c-erbA. Cells containing the reporter gene plus the v-erbA gene also produced much less CAT than controls without a test gene; predictably, this inhibition was unaffected by hormone since the v-erbA gene product does not bind thyroid hormone. The products of both test genes could bind DNA containing the thyroid-responsive element in vitro, and binding was not affected by hormone. Finally, when both test genes were inserted into the same cell, the thyroid hormone receptor could not enhance transcription of the CAT gene in the presence of thyroid hormone. Its enhancing activity was inhibited by the v-erbA protein.

The surprising finding that thyroid hormone receptor binds DNA in the absence of ligand and, further, that it inhibits gene transcription, caused Dr. Evans to question the current paradigm for oncogene action. Scientist have generally accepted the idea that oncogenes act by turning on genes that normally would be off, thus causing cell growth to continue unchecked. Dr. Evans demonstrated that an oncogene can prevent gene transcription and, moreover, that the gene product can compete with the normal regulatory factor for binding. His is the first demonstration of a dominant negative oncogene. How could a negative regulator induce tumors if it turns off genes? One possibility: the repressed gene is necessary for limiting cell growth or differentiation; in the absence of the gene product, the cell would continue to grow indefinitely. Another possibility: the repressed gene inhibits another oncogene which acts synergistically with v-erbA to produce a tumor cell. Regardless of the mechanism, these findings will force scientists to reformulate ideas about tumorigenesis.

For a basic scientist interested in the larger question of how gene expression is regulated, the key observation is that in the absence of ligand, the receptor inhibits gene expression, but when complexed with hormone, it enhances transcription. Although the receptor binds DNA with so-called "zinc finger" structures, nothing is known about how the receptor binds the hormone or how the complex binds DNA. Does the receptor alone prevent RNA polymerase from binding to the promoter? Does the complex stabilize polymerase binding? Little is known about how enhancers actually increase DNA transcription; the erbA system provides an exciting and important model for asking both basic and applied questions about gene regulation.



Jo[illegible]. Jones, Ph.D.

References:

1. Evans, R.M. (1988). The steroid and thyroid hormone receptor superfamily. Science 240, 889-895.

2. Damm, K., C.C. Thompson, and R.M. Evans. (1989). Protein encoded by v-erbA functions as a thyroid-hormone receptor antagonist. Nature 339, 593-597.

3. Umesono, K. and R.M. Evans. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57, 1139-1146.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 14, 1989

From Program Administrator, Genetics Program, NIGMS

Subject Research Highlight: "DNA Supercoiling as a Result of Transcription"
Leroy Liu, Johns Hopkins University (R01 GM 27731),
James C. Wang, Harvard University (R01 GM 24544) and
Rolf Sternglanz, State University of New York, Stony Brook (R01 GM 28220)

To Director, NIGMS
Through: Director, Genetics Program, NIGMS _____
Chief, Molecular and Medical Genetics Section, Genetics Program, NIGMS _____

Enzymes that change the "linking number", or number of helical turns in a stretch of DNA, are abundant in both prokaryotes and eukaryotes. These enzymes, the DNA topoisomerases and DNA gyrases, are known to play important roles in DNA replication, where they act to unwind the two strands of the DNA helix during semiconservative replication. That function, and the role of topoisomerases in forming the superhelical turns that allow DNA winding in nucleosomes, can alone explain why topoisomerases are among the most abundant non-histone chromosomal proteins in eukaryotes.

It has generally been recognized that DNA topoisomerases are also involved in RNA transcription. Usually it has been assumed that transcription responds to changes in DNA linking number -- that supercoiling (under- or over-winding DNA) makes it easier or more difficult for RNA polymerase to open the helix. Recent work of three NIGMS grantees, Leroy Liu (GM 27731), James C. Wang (GM 24544), and Rolf Sternglanz (GM 28220), is turning this relationship between supercoiling and transcription upside down, with broad implications, especially in the areas of gene regulation and genetic recombination.

Maaloe and Kjelgaard pointed out in 1966 that there is a potential topological problem in transcription, resulting from the fact that the DNA strand that serves as template is wound around the axis of the helix. As RNA polymerase tracks this strand (on a fixed helix) the RNA transcript it produces would also be wound around the DNA duplex once every ten base pairs (see Fig. 1). The RNA could be untangled by unwinding it directly, but this would present a problem if the nascent transcript is quickly associated with large complexes. In prokaryotes, where transcription and translation are coupled, ribosomes quickly attach to the nascent message and make it practically immobile, and in eukaryotes spliceosomes might have a similar effect.

Liu and Wang (1) analyzed the problem in detail and suggested another solution. Tangling of the transcript around the DNA duplex would be prevented if the transcribing RNA polymerase were to remain stationary as the template is rotated (see Fig. 2). This would create positive supercoiling (tightening of the DNA helix) ahead of the polymerase and negative supercoiling (loosening) behind. According to the model, these supercoil turns would be quickly relaxed by DNA topoisomerases. However, the model predicts that, in a mutant with a specific defect in the removal of either positive or negative supercoiling, supercoiling of that sign should accumulate in the template DNA as a result of transcription.

This could be seen most easily by using circular plasmid DNA as the template (see Fig. 3).

In E. coli there are two topoisomerases that can relax supercoiling, and both are specific with regard to the sign of the supercoiling. Topoisomerase I relaxes only negative supercoils, and DNA gyrase (also known as topoisomerase II) relaxes only positive supercoils. The Liu and Wang model predicts that removing both enzymes by mutation would have no net effect on the linking number of purified circular plasmid DNA because positive and negative turns generated by transcription would simply cancel. However, if only one enzyme is inactivated, the action of the other enzyme would leave a net accumulation of either positive or negative turns, which the model predicts would be dependent on transcription.

This is exactly what Drs. Liu and Wang and their colleagues have now observed (2). Transcription causes plasmid DNA to accumulate negative supercoiling in topoisomerase I-deficient E. coli mutants and positive supercoiling when DNA gyrase is inhibited by the antibiotic novobiocin.

Analogous observations in yeast have now been made and extended by Giaever and Wang (3) and by Brill and Sternglanz (4). Two enzymes in yeast, topoisomerases I and II, can each relax either positive or negative supercoiling. Topoisomerase 1, encoded by the top1 gene, acts by nicking one strand and passing the other strand through the nick. Topoisomerase II, encoded by top2, makes a transient double strand break, and passes the duplex through the break. (Topoisomerase II can therefore also decatenate intertwined DNA molecules following replication and is essential for this function.)

Brill and Sternglanz examined transcription from the galactose-regulated GAL1 promoter, carried on a plasmid, in Saccharomyces cerevisiae carrying mutations in top1 (null) and top2 (temperature sensitive). In the top1top2 double mutant, transcription at the non-permissive temperature caused the plasmid DNA to become highly negatively supercoiled, containing at least 58 more negative turns per molecule than plasmid DNA from TOP+ cells. The effect was shown to require transcription, not just galactose induction, since it was not observed in a strain containing an inactive (temperature sensitive) RNA polymerase. Furthermore, the magnitude of the effect was proportional to the length of transcript produced.

The effect was not unique to the GAL1 promoter, or to RNA polymerase II transcription. Supercoiling was observed as well with a variety of plasmids containing different promoters, including the RNA polymerase I-transcribed ribosomal RNA promoter and promoters transcribed by the heterologous polymerase of E. coli bacteriophage T7 (which has been shown to be capable, when produced in yeast, of entering the nucleus, recognizing, and transcribing T7 promoters). T7 promoters, especially, are unlikely to share any activators with normal yeast promoters.

Such excessive supercoiling was not observed in TOP+ strains or in a top2 single mutant and was reduced in a top1 single mutant. Apparently, topoisomerase I is primarily responsible for removing transcription-created negative supercoils in

TOP+ strains, and topoisomerase II can substitute at a reduced rate. The fact that negative supercoils are produced in the double mutant implies (a) that the balancing positive supercoils must be removed, in the double mutant, by a third, unknown, topoisomerase (or by residual activity of the temperature-sensitive topoisomerase II); and (b) that the plasmid in the cell is organized into at least two independently supercoiled domains (otherwise the positive and negative turns would cancel on a circular plasmid). Such domains could result, most simply, from multiple convergent or divergent transcripts on the same plasmid. A more interesting, but at this point unnecessary, explanation would be the existence of points of attachment of plasmid DNA to a nuclear matrix.

These results suggest that the problem of tangling of RNA transcripts around their DNA template is avoided by rotating the template. This model points to transcription as an important source of DNA supercoiling, and suggests that the local supercoil density in a region of DNA is influenced by transcription. Local supercoiling could in turn modulate a variety of processes, including gene expression, DNA replication, DNA repair, and genetic recombination. There have been several observations in yeast suggesting that transcriptional activity in a given region may be a significant determinant of recombination frequency nearby. For example, NIGMS grantee G. Shirleen Roeder (Yale University; GM28904) has found that a DNA sequence (HOT1) that stimulates recombination in its vicinity is really a set of sequences that regulates ribosomal RNA transcription, and that transcription by RNA polymerase I stimulates genetic recombination (5). The Liu and Wang model may provide a general rationale for such effects. Finally, these studies increase appreciation of the topological problems presented by linear information transfer from helical DNA and the importance of DNA topoisomerases in preventing the massive tangles that would otherwise result.

Stephen R. Fahnestock

References:

1. Liu, L.F., and Wang, J.C. (1987) Proc. Natl. Acad. Sci. USA 84, 7024-7027.
2. Wu, H.Y., Shyy, S., Wang, J.C., and Liu, L.F. (1988) Cell 53, 433-440.
3. Giaever, G.N., and Wang, J.C. (1988) Cell 55, 849-856.
4. Brill, S.J., and Sternglanz, R. (1988) Cell 54, 403-411.
5. Voelkel-Meiman, K., Keil, R.L., and Roeder, G.S. (1987) Cell 48, 1071-1079.
6. Futcher, B. (1988) Trends in Genetics 4, 271-272.

Figures
(Figs.1 and 2 from ref.6; Fig.3 adapted from ref.2)

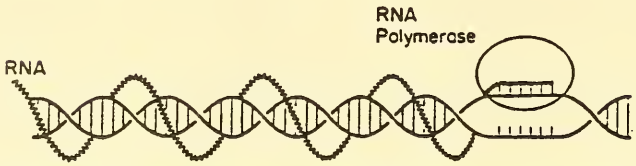


Fig. 1. If RNA polymerase follows the template strand around the axis of the DNA, then the transcript becomes wrapped around the helix once every 10 bp.

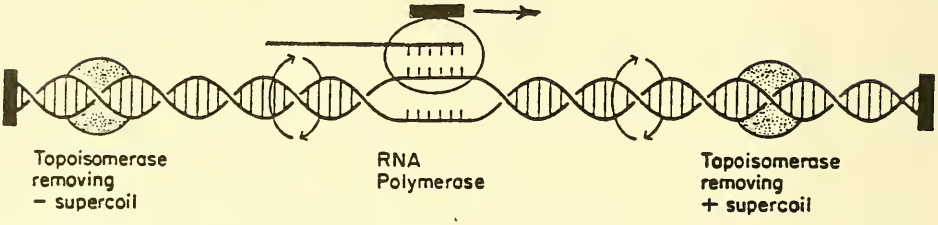


Fig. 2. The transcript does not become entangled if the DNA rotates beneath a fixed polymerase. If the ends of the DNA are fixed, perhaps by the scaffold attachments of adjacent transcription units (black boxes), then rotation of the DNA would generate supercoils. The supercoils could be relaxed as they arose, by topoisomerase swivels. Topoisomerases are known to be specifically associated with transcribed regions in eukaryotes[22]. In the absence of the negative supercoil-relaxing topoisomerase on the left, negative supercoils would accumulate. In the absence of the topoisomerase on the right, positive supercoils would accumulate.

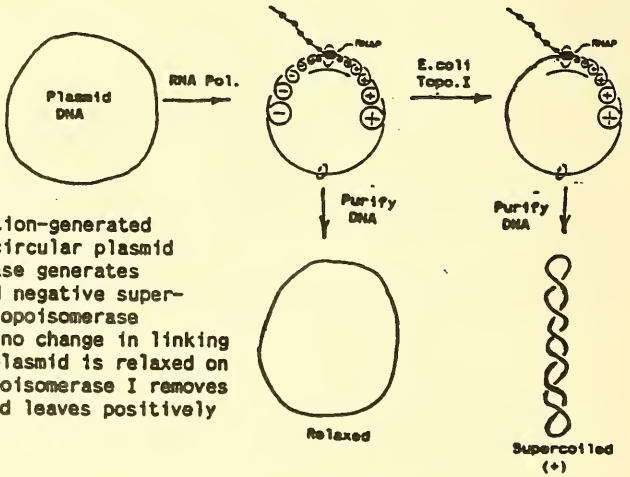


Fig. 3. Transcription-generated supercoiling on circular plasmid DNA. RNA polymerase generates both positive and negative supercoiling. Before topoisomerase action, there is no change in linking number, and the plasmid is relaxed on purification. Topoisomerase I removes negative turns and leaves positively supercoiled DNA.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 25, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Gene Regulation in Tomatoes: A Ripe Area for Research"
R01 GM 33856-04 (Fischer, R.), University of California, Berkeley

To Director, NIGMS
Through: Director, Genetics Program _____

Hormones can have pronounced effects on gene expression. Numerous systems are under study, most notably gene regulation by steroidal hormones. Those studies are yielding exciting information about a family of structurally and functionally related receptors for a number of hormones and vitamins that affect gene activation and cell growth or differentiation (see Reference 1). Many other hormones besides steroids affect gene expression, and a full understanding of how exogenous factors alter gene expression requires study of many different model systems, including plants. One interesting system under investigation by an NIGMS grantee, Dr. Robert L. Fischer, is ethylene-induced tomato ripening. Dr. Fischer is one of a number of molecular biologists who are interested in understanding gene regulation in plants. His work demonstrates the utility of plants for studying universal biological phenomena.

Ethylene profoundly influences many stages of plant development including seed germination, vegetative growth, and fruit development. In fruit, exogenous ethylene greatly hastens ripening whereas ethylene inhibitors greatly retard it. Ripening tomatoes (and other fruit) also make ethylene. Using mature green (MG) tomatoes (full-sized, but not ripe) at identified stages of development, Dr. Fischer and his colleagues at Berkeley confirmed that MG1 (just full grown) and MG2 tomatoes (gel just beginning to form) made no more ethylene than immature tomatoes (half-sized); MG3 fruit (gel formation complete) made slightly higher levels of the hormone, while MG4 tomatoes (internal red color detectable) made very high levels. Molecular analysis of ripening fruit identified at least eight different mRNAs that were more abundant in MG4 than in MG1 tomatoes; these mRNAs could also be detected in MG1 fruit after two hours of ethylene exposure. The mRNAs were isolated and used to make cDNA probes to investigate the expression of four ethylene responsive genes.

At this time, the product of only one of these four "genes" is known, and their names (E4, E8, E17 and J49) are quite arbitrary, reflecting the isolation scheme rather than any special characteristic. All respond to ethylene with a fairly sharp dose response curve, and the response induced by the hormone is inhibited by a competitive inhibitor of ethylene. In many ways, however, they are very different and represent the range of responses that exogenous stimulators can elicit. Two of the "genes" are actually

small gene families with three (E8) or five (J49) members. Three of these genes (or gene families) respond to ethylene in leaves as well as in fruit, whereas the fourth (E8) responds only in fruit. Two, E4 and E8, have a "typical" response to an exogenous inducer such as ethylene: shortly after exposure, the levels of nuclear RNA (the unprocessed RNA transcribed from DNA) and messenger RNA (mRNA, RNA from which the introns have been cleaved) rise. The level of RNA depends both on the length of exposure and the concentration of ethylene. During natural ripening, the increase in the level of both nuclear and mRNA in fruit parallels the production of ethylene.

J49 is different. This gene family responds to ethylene across a fairly narrow dose range, with relatively high levels inhibiting RNA production. During ripening, J49 RNA can be detected before the fruit begins making ethylene, and its level falls when the level of endogenous ethylene peaks. Dr. Fischer suggests that J49 may be more sensitive to very low levels of ethylene as ripening begins, inducing transcription initiation; as ripening progresses, ethylene inhibits transcription. The mechanisms are not known.

E17 is equally complex. There is little E17 RNA (nuclear or messenger) in immature fruit; mature green fruit (MG1 stage) have the same low level of mRNA but high levels of nuclear RNA. Transcription apparently begins before ripening, but mRNA is not produced. Natural ripening or exposure to ethylene have little effect on nuclear RNA levels, which are already high, but cause a significant increase in mRNA. Ethylene's effects are evident after transcription, either stabilizing mRNA or altering RNA processing or transport from the nucleus. The cDNA sequence of E17 was compared with sequences in GenBank; the gene appears to encode a small protein related to the proteinase inhibitors, a diverse class of proteins that accumulate in damaged tissues. This gene product may function to "protect" fruit from damage.

Dr. Fischer's work is only a beginning, but it sets the stage for an exciting inquiry into gene regulation. He has recently succeeded creating a transgenic plant in which he has inserted the E8 gene and demonstrated that it is properly regulated by ethylene, the first time that a hormonally-regulated gene has been experimentally transferred between plants. Ethylene is an interesting example of a hormone which is fairly simple chemically but has very complex biologic effects, affecting both transcription initiation and the production of mature mRNA. It holds great promise as a tool for unraveling some of the complexities of gene regulation.

Plants have historically held an important place in genetics, from Gregor Mendel's peas to Barbara McClintock's maize. The advent of molecular biological techniques gave animal and yeast cell cultures a temporary advantage, but plant systems are rapidly catching up. In addition to their obvious agricultural importance, plants offer a variety of systems for studying many fundamental problems. Maize, a monocot, is widely studied, and dicotyledonous plants such as tomato, soybean, and, of course, tobacco are increasingly being used as experimental systems. In many cases entire plants can be regenerated from tissue culture cells and subsequently propagated

through seeds. Their regenerative capabilities combined with the production of a very large number of offspring makes plants very attractive models for studying genetics and molecular biology.

Joye F. Jones, Ph.D.

References

1. Giguere, V., N. Yang, P. Segui, and R.M. Evans. 1988. Identification of a new class of steroid hormone receptors. Nature 331, 91-94.

2. Lincoln, J.E., S. Cordes, E. Read, and R.L. Fischer. 1987. Regulation of gene expression by ethylene during Lycopersicon esculentum (tomato) fruit development. Proc. Natl. Acad. Sci. USA 84, 2793-2797.

3. Lincoln, J.E. and R.L. Fischer. 1988. Diverse mechanisms for the regulation of ethylene-inducible gene expression. Mol. Gen. Genet. 212, 71-75.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 30, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "What's a Nice Intron Like You Doing in a Gene Like This?"
R01 GM 32422 (Walbot, Virginia), Stanford University

To Director, NIGMS
Through: Acting Director, Genetics Program, NIGMS 3/31/88

A number of years ago, scientists made the startling discovery that in most eukaryotic cells, the coding region of a gene is not continuous, but is interrupted by stretches of non-coding DNA, sometimes many hundreds of nucleotides long. Since then, many scientists have been trying to understand why these regions--called introns--are there. Intuitively, introns would seem to be a waste of energy and resources for a cell. What role, if any, do they play?

Dr. Virginia Walbot and colleagues at Stanford were interested in studying the control of expression of a gene that encodes alcohol dehydrogenase (Adh) in maize. They developed a system for introducing DNA into maize cells growing in tissue culture. Since maize cells normally make Adh, Dr. Walbot inserted a gene for a variant form of the enzyme (called Adh-S) into tissue culture cells and showed that the product of the inserted gene was made in essentially the same amount as that of the native gene (called Adh-F). The scientists were now in a position to study gene regulation in vitro instead of having to work with the whole plant, and further, could work with engineered genes that they introduced into the cells.

The scientists were particularly interested in studying the non-coding regulatory region of the gene called the promoter. This is a stretch of DNA which usually precedes the coding region and to which the enzyme RNA polymerase binds. The polymerase then moves along the gene, making an RNA copy of the DNA, a process known as transcription. Introns in the RNA are subsequently cut out, and the resulting mRNA is translated into the protein gene product. Gene transcription can be measured by determining the amount of the protein gene product being made by the cell. Some gene products, such as Adh, are relatively difficult to measure, and scientists have developed ways to get around the problem. Instead of looking for the native product of the gene, they attach the promoter region from the gene of interest to the coding region of another gene encoding a protein that is easier to detect, frequently an enzyme called chloramphenicol acetyltransferase, or CAT. Thus, researchers can study gene regulation, including gene promoters, using an easily measured end product.

Dr. Walbot and colleagues constructed a gene in which they attached the Adh promoter to Adh cDNA, and transferred this artificial gene into maize cells. (cDNA is the DNA copy of a piece of mRNA; it contains the coding region of the

gene but lacks introns.) Very little Adh encoded by this artificial gene could be detected in the cell containing this gene. The scientists then constructed a gene that contained the Adh promoter linked to the coding region of the CAT gene so that the gene product would be easier to detect. To their surprise, they found that the cells containing the chimeric gene made only minute quantities of CAT, and their controls indicated that the result was not due to a technical problem. What was different about the artificial "genes" compared with the native genes? Introns: the native Adh gene contains nine; the artificial Adh-cDNA gene and the Adh-CAT chimeric gene, none. Could the difference in the levels of gene expression be explained by introns? There were reports in the literature of similar findings in other systems ranging from rabbits and mice to yeast which suggested that introns might have a critical role in controlling gene expression.

In a true tour de force, the scientists constructed a large number of artificial genes. Some were Adh-cDNA genes with one or more introns inserted. Some were Adh-CAT chimeras with Adh introns. In some chimeras, the introns were placed before the promoter, in others after the coding region, and in still others between the promoter and the coding region. A high level of gene expression occurred only when the first (and largest) intron, called intron 1, was inserted between the promoter and the coding region. This was true whether the "reporter" gene encoded CAT, Adh, luciferase or a protein causing resistance to the drug neomycin; whether the promoter was from maize Adh or a cauliflower mosaic virus gene; and whether the intron was from Adh or another maize gene called Bronze. To determine whether it was the specific sequences in the intron that were critical or just the presence of any nucleotides, the scientists replaced intron 1 with an equal length of cDNA (they used cDNA rather than a random DNA sequence so that no "stop" codons would be present to terminate transcription or translation). This cDNA "intron" was placed between the promoter and the gene's coding region; no protein product was made.

Since the scientists were measuring the protein gene product, the effect of the introns could be on either transcription or translation. In some of their experiments, therefore, the scientists measured mRNA production rather than protein production. Without introns, essentially no mRNA was produced, which is consistent with the hypothesis that introns affect transcription. However, the scientists measured mRNA which has already been processed, rather than RNA as it was made by the polymerase reading the DNA. An alternative interpretation of their findings is that introns are necessary not for transcription but for some other step such as transporting RNA from the nucleus into the cytoplasm or processing it into mRNA.

We still don't know what introns do. Dr. Walbot suggests that introns may affect RNA conformation which in turn somehow affects gene transcription. Certainly introns are not universally required, because the expression of some genes in plants, animals, and viruses are unaffected by the removal of introns, and prokaryotic (bacterial) genes rarely have introns. Futhermore,

many scientists use the CAT gene--without introns--as a "reporter" gene very successfully, although it is possible that adding an intron might make those constructs work even better. Clearly, however, for many genes, introns are critical for efficient production of mRNA.

Joyce F. Jones, Ph.D.

Reference:

Callis, J., M. Fromm, and V. Walbot. 1987. Introns increase gene expression in cultured maize cells. Genes and Development 1:1183-1200.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 2, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Exon Shuffling" R01 GM36714 (Dwight H. Hall), Georgia Institute of Technology, and R01 GM37746 (David A. Shub), State University of New York at Albany

To

Director, NIGMS
Through: Director, Genetics Program, NIGMS 11-7-89
Chief, Molecular and Medical Genetics Section, Genetics Program, NIGMS 11-2-89

The genes of higher organisms are generally interrupted by extra DNA sequences, called introns, that are transcribed into RNA along with the coding sequences of genes, but must be removed before the RNA transcripts can function as message RNA (or as ribosomal or transfer RNA). Nearly all human genes are interrupted by at least one, and often many, introns. An extreme example is the dystrophin gene, mutations in which are responsible for Duchenne muscular dystrophy. This gene encodes a large protein of 3685 amino acid residues which functions in muscle development. Expression of this gene begins with transcription, forming an RNA about two million nucleotides long, which contains over 100 introns, the longest of which is more than 200,000 nucleotides long (more than 20 times the length of the entire coding sequence). Before this RNA can direct the synthesis of dystrophin all of these introns must be precisely excised by an apparatus called the splicecsome, which is found in the nucleus. Thus, in this huge gene, comprising about 0.1% of the entire human genome (the full complement of chromosomes), only 0.2% of the DNA actually codes for its protein product. The other 99.8% encodes introns that are removed and discarded during expression!

Since their discovery over 10 years ago, the function of introns has remained a mystery. It appears that introns were present in the genes of the ancient ancestor of all living organisms, but have been largely removed from the genes of bacteria and lower eukaryotes. Why have they been retained by higher eukaryotes? One possibility is that they are simply parasitic and difficult enough to remove that unless they impose a significant burden (as in the case of organisms like bacteria that must keep their genomes small to allow for rapid growth), it is more trouble to remove them than it is worth. This hypothesis is not very satisfying, however, because it seems to excuse a great deal of inefficiency in basic life processes. Therefore, there have been several attempts to identify an advantage that introns might confer.

One of the earliest of such attempts was a proposal by Walter Gilbert (1). Dr. Gilbert considered the evolution of intron-containing genes, and proposed that introns would provide useful targets for recombination that would allow coding pieces of genes (the exons) to be combined in various assortments. If recombination occurred in introns and the resulting hybrid introns could be excised correctly, then the coding sequences could be joined precisely regardless

of the exact site of recombination. This hypothetical process, called "exon shuffling," would greatly increase the frequency of productive intergenic recombination events and thereby facilitate evolution. The advantage would not be apparent to the vast majority of individual organisms carrying the burden of introns, but over the long haul of evolution the most adaptable species would be those that carried introns.

Such a mechanism for generating new proteins would work best if individual blocks of coding sequence were more or less self-contained, or able to fold into stable structures that could be combined with each other with a high probability of generating a stable combined structure. According to this view, proteins may be constructed from structural modules that can be assorted to create novel molecules with relatively high probability of forming useful structures. The model therefore predicted that individual exons would tend to encode structural domains of proteins. This was subsequently found to be true in many cases: introns often tend to occur at domain boundaries. Furthermore, there are several examples of homologous domains encoded by exons in different genes, with analogous intron boundaries. However, there are many exceptions as well.

Biological molecules often give clues to their history, and comparisons across species have shed much light on the evolutionary history of living organisms. Such evidence is always indirect, however. Now the work of two NIGMS grantees, Dr. Dwight Hall and Dr. David Shub (2), has for the first time demonstrated directly that the process of exon shuffling can occur. The experimental system they used was bacteriophage T4, a virus that infects the bacterium Escherichia coli. The discovery several years ago of introns in several T4 genes was surprising to most scientists because it had previously been thought that introns were limited to eukaryotes. Not only did this discovery increase appreciation of the generality of introns, but it also provided an extremely powerful system in which genetics could be applied to study the function and mechanism of excision of introns.

Drs. Hall and Shub examined spontaneous deletion mutations affecting two T4 genes, td and nrdB, which encode, respectively, thymidylate synthase and a subunit of ribonucleoside diphosphate reductase. Both of these genes contain introns of a type known as Group I, with sequence homology to each other and to a group of introns in eukaryotes. These introns are self-splicing; that is, the intron RNA catalyzes its own removal and the splicing of the adjoining ends of the message RNA. The td and nrdB genes are located close together on the T4 genome, with no essential genes between them, and are transcribed in the same direction. Drs. Hall and Shub surmised that recombination between homologous regions of the two introns might generate a hybrid intron that could still carry out splicing and would effectively join the upstream exon of td with the downstream exon of nrdB.

A mutant phage [designated del(63-32)] was identified in which just such a recombination event had occurred spontaneously, resulting in the loss of both td and nrdB function. DNA sequencing revealed that a recombination event had occurred between conserved regions of the two introns, resulting in the deletion of DNA sequences between them. Drs. Hall and Shub were then in a position to ask whether the resulting fused gene is transcribed, and whether the resulting mRNA is spliced. The answer to both questions was "yes". The resulting hybrid

mRNA was sequenced by extending a primer with reverse transcriptase, a powerful method for sequencing RNA from complex mixtures. The sequence revealed that extensive splicing had taken place, and while the 5' splice site was identical to the normal td site, the 3' splice site was shifted one nucleotide from the normal nrdB site. This shift could be rationalized based on the sequence of the splice site determining region of the intron.

Even though, in this case, no functional hybrid protein was produced, these observations provide important direct evidence that exon shuffling can occur, and that hybrid introns generated in the process can be excised. There is now little doubt that exon shuffling to facilitate the generation of genes for novel proteins is a potential evolutionary benefit of introns. Whether there are additional advantages of more immediate value to the organism remains to be determined.

S. Fahnestock

Stephen R. Fahnestock, Ph.D.

References:

1. Gilbert, W. (1978). Nature 271, 501.

2. Hall, D.H. and Shub, D.A. (1989). Exon shuffling by recombination between self-splicing introns of bacteriophage T4. Nature 340, 574-576.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 2, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Infectious Introns" R01 GM 39422 (Marlene Belfort), State University of New York, Albany; R01 GM22525 (Ronald A. Butow), University of Texas Southwestern Medical Center; R01 GM31480 (Philip S. Perlman), Ohio State
To University; R01 GM31577 (Volker M. Vogt), Cornell University

Director, NIGMS
Through: Director, Genetics Program, NIGMS 11-7-89
Chief, Molecular and Medical Genetics Section, Genetics Program, NIGMS 11-2-89

The coding sequences of most genes of higher eukaryotes are interrupted by introns, non-coding sequences that are excised from message RNA before it can be translated. In lower eukaryotes, nuclear protein-coding genes are mainly free of introns, but introns are common in genes for transfer RNA, and in genes in mitochondria. Introns have also been found in prokaryotic systems, but so far only in bacteriophage genes. The origin of introns has been a subject of debate among scientists since their discovery more than ten years ago. One possibility is that introns were inserted into genes of higher eukaryotes relatively recently, and were not present in the genes of the common ancestor of prokaryotes and eukaryotes. That seems unlikely in view of their presence in mitochondria, which are now generally believed to be derived from ancient invading prokaryotes. The alternative view is that introns were present in the most ancient genes and have since been lost from genes of bacteria and lower eukaryotes.

A consensus seems to have formed around this latter view, which holds that introns were present in the genes of the earliest ancestor of both prokaryotes and eukaryotes. Thus, studies of common intron positions in members of multigene families in organisms of many different species have been interpreted as evidence that some introns must have been present before the divergence of members of the multigene family. However, the same studies have identified introns that appear to have been inserted later. For example, introns in tubulin and actin genes are located at different positions in different species. The positions of some are common to many species, and therefore probably reflect ancient presence. However, others appear at so many different positions that it is very unlikely that the current distribution could have originated solely by the loss of some introns from an ancestral pattern that included all of the observed positions. For example, two of the observed positions are one nucleotide apart. If both of these introns had ever been present in the same gene, the exon between them would have been one nucleotide long, a very unlikely situation. Therefore, at least one of them must have been inserted after the divergence of the two species.

Such arguments provide circumstantial evidence that introns can move. The work of several NIGMS grantees has now provided direct evidence that introns are mobile.

Dr. Marlene Belfort, who discovered the first introns in prokaryotic genes, those of bacteriophage T4, five years ago, has found convincing evidence that two of them are mobile. Dr. Belfort and her coworkers (1) asked whether introns from T4 genes could be transferred to the closely related phage T2, which contains no known introns. They inserted cloned fragments of T4 DNA with intron-containing genes into E. coli, then infected those cells with T2. When the intron-containing T4 sunY gene was present, 55.8% of the progeny T2 phage had acquired the intron in their homologous sunY genes. The intron in the T4 td gene was transferred to 18.1% of the T2 progeny, but the T4 nrdB intron was not transferred at all.

One possible mechanism for intron insertion would involve a reversal of the splicing reaction at the RNA level, reverse transcriptase synthesis of a DNA copy of the RNA with inserted intron, and insertion into the chromosome by homologous recombination. Such a process does appear to have been responsible for some events in eukaryotes. In this case, however, it could be excluded by the observation that the frequency of transfer of the td intron was not decreased by a mutation in the intron that interfered with splicing. Therefore, it appeared that transfer was occurring at the DNA level.

All three known T4 introns contain open reading frames (ORFs) that code for proteins expressed only late in infection. Dr. Belfort found that transfer of the td and sunY introns depended on expression of their ORFs. The required intact ORF could be provided either in cis (on the transferred intron) or trans (elsewhere in the cell), and by increasing the level of expression of the sunY ORF, transfer could be increased to 85% of the progeny T2 phage. The activity of the ORF-encoded protein is therefore rate limiting for the transfer process. The proteins encoded by the td and sunY ORFs were found to be endonucleases that specifically cut phage DNA at the sequences into which the intron inserted. Thus, the introns appear to be inserted by a mechanism that involves site-specific recombination initiated by sequence-specific cleavage at the insertion site, catalyzed by the intron-encoded enzyme.

Their mobility makes the phage T4 introns look very much like molecular parasites, and their absence from the closely related phage T2 suggests that they are relatively recent arrivals. However, a closer look at the expression of the intron ORFs suggests a long association with the phage. The intron ORFs are expressed only late in infection, a strategy that can be rationalized as a means of avoiding transfer of the intron to the host DNA which is destined to be degraded. This is true even though they are contained within genes that are expressed earlier. How is expression of the ORF avoided even though the intron is transcribed earlier in infection? The answer appears to be that the ribosome initiation signals required for translation of the ORF are sequestered by RNA secondary structure when the entire intron-containing gene is transcribed. Late in infection the ORF is transcribed from a promoter within the intron, and this transcript lacks upstream sequences involved in sequestering the initiation site of the longer early message. This late promoter is under control of the T4 gene 55 product, a protein that associates with E. coli RNA polymerase to allow it to favor late T4 promoters. Thus, the intron appears to be well integrated into the physiology of the phage, and behaves more like a native than a foreigner. It has also lost whatever foreign accent it may have arrived with; codon usage

in the ORFs follows the highly characteristic T4 pattern. Here the line between host and parasite is blurred, probably by a long association, but the amount of time involved is difficult to estimate.

Intron mobility has also been demonstrated in other systems. The first demonstration of intron mobility came from studies of a self-splicing intron in a yeast mitochondrial rRNA gene. This intron, found in some yeast strains but not in others, appears to be transferred by a mechanism similar to that described for the T4 introns, also involving a sequence-specific endonuclease encoded by the intron. Similar mobility of several other yeast mitochondrial introns has since been demonstrated. Drs. Ronald Butow, Philip Perlman, and their coworkers (2) have recently found that one such mobile mitochondrial intron (intron 4 of the gene encoding subunit I of cytochrome c oxidase, or aI4α, encodes a protein with two functions. Like the ORF products of the T4 introns, it is a site-specific endonuclease required for intron mobility, but it also functions as a maturase that participates in the splicing of the intron at the RNA level (by stabilizing the folding of the catalytic RNA). ORFs in other mitochondrial introns encode proteins that function either as DNA endonucleases or maturases, but none had previously been shown to have both functions.

Mobility of a nuclear intron has been demonstrated by another NIGMS grantee, Dr. Volker Vogt (3). In the slime mold Physarum polycephalum, genes for ribosomal RNA are located on extrachromosomal DNA molecules, which is unusual. Also unusual is the presence of an intron. Introns are generally not present in nuclear rRNA genes in lower eukaryotes, but are limited to mitochondrial and chloroplast rRNA genes. Dr. Vogt has found that this nuclear rRNA intron is mobile, and can be transferred from an intron-containing rDNA molecule to an intron-less one when two nuclei fuse during mating. As with other mobile introns, this transfer depends on the activity of a sequence-specific endonuclease encoded by the intron.

The existence of "infectious" introns supports the idea that at least some introns could have been inserted into genes relatively recently. Furthermore, the sequence-specific nature of the insertion process, depending on the specificity of an intron-encoded nuclease, provides a possible explanation for the occurrence of introns at homologous sites in members of a gene family even though they may have arrived after divergence of the genes. This alternative makes less compelling the argument that introns must have predated the divergence of the gene family if they are located at the same site in several genes of the family. Infectious introns therefore confound attempts to understand the evolution of introns and their "hosts" by constructing lineages based on divergence. But at the same time they inspire respect for the elegance and economy of these simplest of parasites (4).

S. Fahnestock

Stephen R. Fahnestock, Ph.D.

References:

1. Quirk, S.M., Bell-Pedersen, D., and Belfort, M. (1989). Intron mobility in the T-even phages: High frequency inheritance of group I introns promoted by intron open reading frames. Cell 56, 455-465.

2. Wenzlau, J.M., Saldanha, R.J., Butow, R.A., and Perlman, P.S. (1989). A latent intron-encoded maturase is also an endonuclease needed for intron mobility. Cell 56, 421-430.

3. Muscarella, D.E., and Vogt, V.M. (1989). A mobile group I intron in the nuclear rDNA of Physarum polycephalum. Cell 56, 443-454.

4. Lambowitz, A.M. (1989). Infectious Introns. Cell 56, 323-326.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 24, 1987

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Are Jumping Genes Spliced When They No Longer Jump?"
R01 GM 32528-05 (Susan R. Wessler), University of Georgia

To Director, NIGMS
Through: Director, Genetics Program, NIGMS
Deputy Director, Genetics Program, NIGMS

Dr. Barbara McClintock was the first to carry out a systematic genetic analysis of the unstable pigment mutations in the kernels of corn. She presented convincing evidence (1) demonstrating that the instability of these mutations was due to the insertion and deletion of DNA elements, called controlling elements. Controlling elements and other transposable elements have now been found in many organisms and exhibit a variety of molecular structures and behaviors. During the past five years, the cloning and sequencing of a number of transposable elements from different organisms have confirmed Dr. McClintock's predictions at the molecular level. Due to Dr. McClintock's initial insights, much of the most informative genetic research on transposable elements is still being carried out in the well characterized maize system.

The first controlling elements described by Dr. McClintock, Activator (Ac) and Dissociator (Ds), were associated with chromosome breakage. The Ac element transposes autonomously whereas the Ds element requires the presence of the Ac for both transposition and chromosome breakage. Both of these genetic elements were cloned and their DNA sequences determined by a number of laboratories, including that of NIGMS grantee Dr. Nina Fedoroff at the Carnegie Institution in Baltimore (GM 34296) (2). The DNA sequences revealed that the Ds element is identical to Ac except that a portion of the transposase gene (the gene that codes for an enzyme essential to movement of the element) is deleted in Ds (figure 1). This explains why the Ds element is not able to transpose in the absence of Ac. Upon insertion of either element into a gene, an 8 bp sequence of host DNA is duplicated at each end of the element. When the element is excised, the duplicated sequence is left behind, sometimes producing profound effects on gene expression. These leftover duplications have been called "dinosaur tracks" since they mark the position of excision events that may have occurred thousands of years ago. The impact of the insertion/excision event on gene expression depends upon whether it occurs within the intron or the exon of the gene.

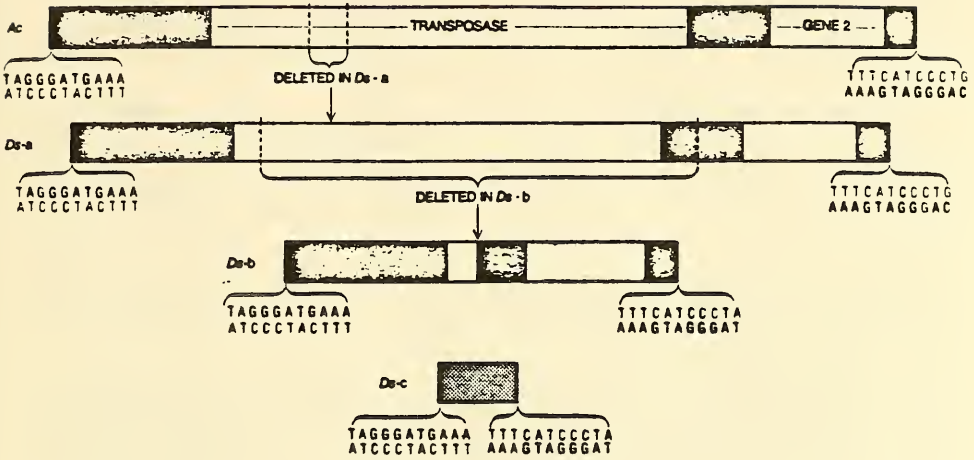


Figure 1: The Ac and several Ds elements of maize. The Ac element is about 4,500 nucleotides long; it has two coding sequences (gray), three noncoding sequences (dark gray), and imperfect inverted terminal repeats (black). The first Ds element isolated, Ds-a, is almost identical to Ac, except that a 194-nucleotide segment (broken lines) of Ac's larger gene has been deleted. That is enough to disable the transposase gene, and it accounts for the inability of Ds to transpose itself. In Ds-b a much longer central segment of Ac has been deleted. Ds-c is very short and retains only the inverted terminal repeats of Ac, suggesting that any piece of DNA bracketed by these terminal sequences can be recognized and transposed by the transposition enzymes and can therefore act as a Ds element.

The waxy (wx) locus of maize has numerous mutations caused by the insertion of Ac or Ds. The waxy locus encodes a starch granule-bound glucosyl transferase involved in starch biosynthesis. Mutant kernels appear waxy, which accounts for the locus name. The presence of Ds in waxy allowed Dr. Susan Wessler, while a postdoctoral fellow in Dr. Fedoroff's laboratory, to identify and isolate the waxy gene (3).

After establishing her own laboratory at the University of Georgia, Dr. Wessler continued research on maize genetics with an investigation of the mechanism of Ds excision in the presence of Ac. In several unstable alleles of waxy she demonstrated that the elements were inserted into both introns and

exons. By sequencing the stretches of DNA where Ds insertion and excision had occurred, Dr. Wessler found that reduced levels of enzymatic activity in the mutants correlated with the presence of "dinosaur tracks" in the exons. While the level of mRNA produced in the mutants was the same as in the wild type, the presence of the duplicated sequences in the exon resulted in a peptide with altered amino acid composition and reduced enzymatic activity.

Dr. Wessler's laboratory has most recently been studying the effect of Ds insertion/deletion in genes when Ac is not present (4). One would expect that because of the lack of an Ac element, the Ds element would not be excised. As a result, Ds would either appear in the mRNA if inserted in the exon of the gene (causing significant disruption of gene expression), or it would be spliced out if inserted into an intron. Surprisingly, in one mutant, wx-m9, where the Ds element was inserted into an exon, gel analysis of the Waxy protein product revealed that a protein with a size of the wild type was still produced, but that it had reduced enzymatic activity. This surprising result suggested that a novel mechanism might be operating whereby the Ds sequences were removed from the transcript.

In order to determine how the Ds elements were excised in wx-m9, Dr. Wessler characterized the mRNAs by Northern blot hybridization. wx-m9 transcripts were compared with those of another mutant, wxB4, which produces no active enzyme presumably because of a Ds present in an exon. Surprisingly, when the blots were probed with waxy DNA, both mutants contained RNA transcripts of 2.4 kb, the same size as the wild type transcript. Reprobing the blots with Ds DNA showed that neither transcript had retained the Ds sequences, but the wx-m9 mutant had a minor transcript of 6.7 kb that was homologous with both Ds and waxy sequences. These data suggested that the 6.7 kb unspliced transcript contained the entire Ds element and the waxy mRNA. The greater abundance of the 2.4 kb mRNA as compared to the 6.7 kb mRNA indicated that the Ds element was being efficiently removed from the unspliced transcript.

The most obvious explanation of how Ds could be excised in the absence of Ac was that intron-like splicing of the element had occurred. Splicing outside of introns had never been reported before, but Dr. Wessler reasoned that it was theoretically possible if the correct DNA consensus sequence for an acceptor and donor site could be found in the DNA. To determine whether splicing had occurred, Dr. Wessler prepared and sequenced cDNAs made from wx-m9 mRNAs and clones of genomic waxy locus DNA from the mutants. Donor and acceptor sites that could have been used in splicing out of the Ds sequences were found in the terminal Ds sequences and the bordering waxy sequences (figure 2). The wx-m9 cDNA that was sequenced did not maintain the correct reading frame for protein translation and so it is unlikely that wild type sized protein detected in wx-m9 arose from mRNA spliced at those particular sites. However, alternative donor sites were found in the Ds sequence. Splicing at these sites should result in a full length protein, although with a slightly altered amino acid sequence. This correlated well with the fact that wx-m9 has reduced enzymatic activity. Dr. Wessler expects that she will be able to find an mRNA spliced at either of the two sites as she sequences more of the cDNAs from wx-m9.

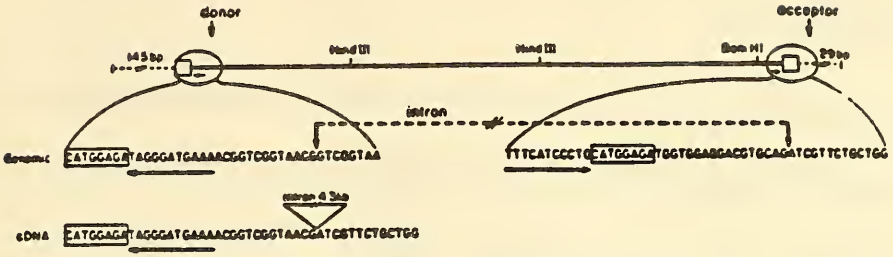


Figure 2: The DNA sequences of genomic and cDNA subclones from wx-m9 reveal how the Ds element is spliced from waxy RNA. The heavy line at the top represents the 4.3 kb Ds element and the position of the element within exon 10. The regions within the ellipses have been expanded to show the genomic sequence at the waxy/Ds junctions. The heavy arrows indicate the 11 bp inverted repeat at the Ds termini. The boxed region and sequences are the 8 bp direct repeat of waxy DNA generated upon Ds insertion. The position of the new intron, as determined by comparing the genomic and cDNA sequences, is shown.

The reason for the evolution of a splicing mechanism to eliminate the Ds element seems to be to restore enzymatic activity of a gene harboring an insertion. Most waxy alleles are like the wxB4 where the splicing throws the transcript out of phase and results in a nonfunctional protein. Similar cases of splicing of nonautonomous elements have now been reported for other genes in maize. In each case the specifics of the splicing are different, but the end result is partial restoration of enzymatic activity. Dr. Wessler expects that there may be many cryptic Ds elements that are efficiently spliced from the transcript and result in no detectable mutant phenotype. She plans to screen cDNAs from total maize mRNA against total genomic maize libraries in an effort to detect cryptic mutants. These experiments will provide a more realistic estimate of the number of Ds elements in the genome and a measure of their true impact on gene expression.

Jane L. Peterson, Ph.D.

References

1. McClintock, B. (1948). Metabolic Loci in Maize. Carnegie Institution of Washington Year Book 47: 155-169.

2. Fedoroff, N., Wessler, S. and Shure, M. (1983). Isolation of the Transposable Maize Controlling Elements Ac and Ds. Cell 35: 235.

Pohlman, R., Fedoroff, N. and Messing, J. (1984). The Nucleotide Sequence of the Maize Controlling Element Activator. Cell 37: 635-643.

3. Shure, M., Wessler, S. and Fedoroff, N. (1983). Molecular Identification and Isolation of the waxy Locus in Maize. Cell 35: 225-233.

4. Wessler, S., Baron, G. and Varagona, M. (1987). The Maize Transposable Element is Spliced from RNA. Science 238: 916-918.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date May 31, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Ubiquitin as a Molecular Chaperone,"
Alexander J. Varshavsky, Massachusetts Institute of Technology (GM 31530)

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 6-2-89

Much of the biology of ubiquitin, including its gene structure and degradative functions, was worked out nearly ten years ago and did not suggest a particularly novel role for the protein in cellular metabolism. However, in a series of important papers over the last three years, Dr. Alexander Varshavsky's laboratory has reported unexpected and diverse functions for this highly conserved cellular protein (see "A New Connection Between DNA Repair and Ubiquitin," June 28, 1988; "Selective Protein Degradation: The N-end Rule," December 30, 1986). In their latest report (1), Dr. Varshavsky and his colleagues have described yet another novel function for ubiquitin: it serves as a covalently bound "chaperone" for ribosomal protein during ribosome biogenesis.

Ubiquitin is encoded by two different, but related, DNA sequences. Most comes from transcription of ubiquitin genes that are arranged tandemly. The polyubiquitin protein product is cleaved to produce mature ubiquitin. Mature ubiquitin synthesized in this way tags proteins destined for degradation. Ubiquitin first binds to a ubiquitin carrier protein, and the complex in turn binds to the target protein. Ubiquitin-containing proteins are also formed by translation of mRNAs from natural gene fusions between ubiquitin and another unrelated gene. Interestingly, the unrelated "tail" polypeptides in these fusion proteins are, like the ubiquitin sequence, highly conserved between species.

In yeast, three different fusion proteins, UBI1, UBI2, and UBI3, are synthesized. The UBI1 and UBI2 proteins contain identical 52-residue tails, whereas the UBI3 protein contains a unique 76-residue tail. However, all three tails are very basic and contain putative DNA-binding domains.

Deletion studies had earlier indicated that the loss of any one of the three genes was not lethal, although the ubi3 deletion severely affected growth. Double mutants of ubi1 and ubi2, on the other hand, were lethal. Since the mRNA and the cellular ubiquitin levels were not markedly decreased in any of the deletion mutants, it seemed that it was not a lack of ubiquitin but rather the loss of the tail polypeptide that caused the deleterious effect.

In order to study the role of the tail polypeptide in cellular metabolism, plasmids containing either the tail or the ubiquitin DNA sequence were constructed. When these were introduced into the mutant cells in which both

the ubi1-2 and ubi3 genes were deleted, the expression of the tail polypeptide complemented the mutation; expression of the ubiquitin protein alone did not complement the mutation. These results confirmed that the tail portion of the protein was required for cellular function and that its fusion to ubiquitin was not essential for this function.

While examining the RNAs of the ubi1-3 mutants electrophoretically, Dr. Varshavsky's laboratory noted that there was an unusually low level of 18S rRNA in the ubi3 mutants. The 18S rRNA resides in the small (40S) ribosomal subunit, and, not surprisingly, sucrose gradient analysis of ribosomes from the ubi3 deletion mutants showed that they had greatly reduced levels of 40S subunits. In pulse-chase experiments, the 20S rRNA (the precursor RNA for 18S rRNA) in ubi3 deletion mutants was very slowly converted to 18S rRNA, and most of it was degraded. Thus, the ubi3 deletion mutants appeared to cause a specific defect in the rRNA processing.

To determine if the deletion of the tail polypeptide sequences was responsible for the rRNA processing defect, the investigators needed to demonstrate that the tail polypeptide could be found in a site where rRNA processing occurs. Two cellular sites are known to be involved in rRNA processing: the snRNPs (small nuclear ribonuclear protein complexes) and the site in the ribosome that recognizes the 20S pre-rRNA. A fused gene was constructed between ubi3 and the myc gene, a cellular oncogene which was chosen because anti-MYC antibodies could be used to detect the location of the gene product. When the ubi3-myc gene was expressed in cells containing the ubi3 deletion, it successfully complemented the ubi3 deletion. Immunoblots using the anti-MYC antibody revealed that the UBI3-MYC protein had become incorporated into the 40S ribosomal subunit. Dr. Varshavsky concluded that the UBI3 protein is a previously undetected ribosomal protein that is involved in rRNA maturation.

Similar analyses of the double ubi1-2 deletion mutants revealed that the tail polypeptide of UB1 and UB2 is a 60S ribosomal protein. Its role, however, is in ribosome assembly and turnover, rather than in rRNA processing.

Further evidence that UB1, UB2, and UB3 are truly ribosomal proteins came from a re-examination of the previously determined DNA sequences found upstream of the ubi1-3 genes. Striking similarities were found between these sequences and the consensus sequence for upstream activation sites (UAS) of genes encoding other ribosomal proteins. In addition, yeast ribosomal genes and the ubi1-3 genes contain introns, a rarity in yeast genes.

Dr. Varshavsky's results have been confirmed in rat cells by work from Dr. Martin Rechsteiner's (GM 37009) laboratory (2). Their results were very similar to Dr. Varshavsky's, and they were able to demonstrate that the larger of the two proteins (UBI3) is a known rat ribosomal protein, S27a. Both groups are now interested in studying the role ubiquitin plays in chaperoning these ribosomal proteins during ribosomal biogenesis. As mentioned above, the association of branched ubiquitin (ubiquitin that is covalently bound to a carrier protein) with a protein usually marks that protein for degradation. Ironically, a possible role for the linear ubiquitin in the UBI1-3 proteins may be to protect the tail polypeptides from degradation long enough for them

to become incorporated into the ribosome. Alternatively, the ubiquitin may increase the rate of transport or assembly of the tail polypeptide into the ribosome. Experiments to distinguish the role of ubiquitin in ribosome assembly are now underway.

Other examples of molecular chaperones have been described, but the association between the two molecules in such cases has been transient and noncovalent. The ubiquitin model is the first in which the chaperoning function is inherent in the DNA sequence. This will make analysis of the protein's processing even more interesting. Ubiquitin continues to surprise Dr. Varshavsky (and the rest of us) by its diversity, and he plans to continue searching for other ways that ubiquitin regulates cell function.

Jane L. Peterson, Ph.D.

References:

1. Finley, D., Bartel, B., and Varshavsky, A. (1989). The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. Nature 338: 394-401.

2. Redman, K., and Rechsteiner, M. (1989). Identification of the long ubiquitin extension as ribosomal protein S27a. Nature 338: 438-440.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 21, 1987

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Unusual Structure of Nematode Mitochrondrial Nucleic Acids" R01 GM 18375 (Wolstenholme, D.) University of Utah

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 8·27·87
Chief, Molecular & Medical Genetics Section, Genetics Program, NIGMS

Mitochondria, the sub-cellular organelles in which the energy-generating reactions of the cell (electron transport and oxidative phosphorylation) occur, are composed of a large number of proteins and other macromolecules. While most of these gene products are encoded in the nuclear genome of the cell, a few mitochondrial proteins and several RNAs required for their synthesis are encoded by DNA that is present in the mitochondrion itself. Analysis of the mitochondrial DNA from a variety of metazoan organisms has shown surprising uniformity in the genetic structure and organization of mitochondrial DNA. Recent work from the laboratory of Dr. David Wolstenholme (GM 18375) has, in contrast, indicated that the mitochondrial DNA of nematode worms has a number of unusual and interesting features which distinguish it from all other known mitochondrial DNAs.

The complete sequence of mitochondrial DNA has been determined for five different organisms, three mammals (human, cow, and mouse), one amphibian (Xenopus laevis) and one insect (Drosophila yakuba). Incomplete sequences are available for a variety of other mammalian and arthropod species (as well as from some plants). Comparison of the sequences has indicated that all animal mitochondrial DNAs contain the same genes. These include the genes for thirteen proteins: cytochrome b, subunits I, II, and III of cytochrome c oxidase, subunits 6 and 8 of ATPase and six subunits of the respiratory chain NADH oxidase; the product of a thirteenth open reading frame, URF6, has not been identified. In addition, each mitochondrial DNA encodes two rRNA molecules and 22 tRNAs. The gene order is the same in the vertebrate mitochondrial DNAs but is somewhat different in the insect mitochondrial DNAs.

A striking feature of mitochondrial DNA is that the genetic code used differs in some respects from that used by almost all other genomes. For example, TGA codes for the termination of protein synthesis in both bacteria and in eucaryotic nuclear messenger RNAs, but codes for tryptophan in the mitochondrial systems. In mitochondrial DNA, AUA specifies methionine instead of isoleucine. In some vertebrates, AGA and AGG, which are arginine codons in the standard genetic code are either rare termination codons or are not used,

while in Drosophila, AGA (but not AGG) is a serine codon. Finally, AT(A, G, T, C) are used as translation initiation codons for some mitochonorial genes, but are not known to be used this way in any other system.

During an analysis of the mitochondrial DNA of several species of Drosophila Dr. David Wolstenholme and co-workers obtained some data that raised the question of whether certain sequences they believed were unusual coding signals in Drosophila mitochondrial DNA were restricted to these insects or were used universally in invertebrates. To answer this question, Dr. Wolstenholme determined the sequence of the mitochondrial DNA of Ascaris suum, a nematode worm which is an intestinal parasite of pigs. However, rather than providing a simple answer to the question that prompted it, this work demonstrated that the Ascaris mitochondrial DNA differed in several ways from the other mitochondrial DNAs whose sequence is known. The most interesting of these differences concerns the sequences which appear to encode the Ascaris analogues of mitochondrial tRNAs.

As noted above, all other known mitochondrial DNAs code for 22 tRNA species which, given the unique nature of the mitochondrial genetic code, seem to be sufficient for carrying out mitochondrial protein synthesis. These tRNAs appear to be similar in structure to procaryotic tRNAs and to the nuclear-encoded tRNAs of eucaryotes. Most importantly, they fit the clover-leaf structure of these other molecules (see Figure 1; also see highlight of October 10, 1986; "Structure-Function Studies of tRNA"), although there are some differences in regard to the frequency of occurence of nucleotides which are highly conserved in nuclear-encoded tRNAs. In typical mitochondrial DNA, the tRNA coding sequences are located, either singly or in clusters of two to six, between protein and rRNA genes.

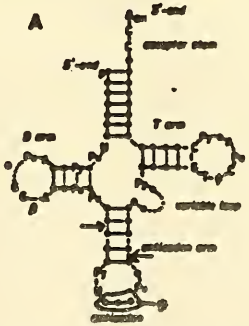


Figure 1. Secondary structure of tRNA, known as the cloverleaf model, showing complimentary base pairing. The letters identify nucleotides which are conserved in all known tRNAs(A, adenosine; G, guanosine; U, uridine; C, cytidine; T, ribothymidine; ψ, pseudouridine; Pu, purine; Py, pyrimidine). Other nucleotide positions are shown as open circles. Thenames given to the various regions of the molecule are indicated (*, anticodon).

In the Ascaris mitochondrial DNA, there are similar spacer regions between the protein and rRNA genes, comparable in size to the tRNA-encoding spacers of other mitochondrial DNAs. However, using computer programs for prediction of RNA secondary structure, transcripts of the Ascaris spacer sequences cannot be folded into secondary structures characteristic of known tRNAs. Yet, all but one of these regions do include sequences which can be folded into structures that have certain features typical of tRNAs (see Figure 2). The important differences include the replacement of the TψC arm and the variable loop with a single loop of from four to twelve nucleotides. One other structure that resembles the $tRNA^{ser}$ of vertebrate and Drosophila mitochondrial DNAs is also found in Ascaris, making a total of 22 tRNA-like sequences (the same as the number of tRNAs found in other mitochondrial DNAs). Several considerations make it seem plausible that these sequences do encode a novel sort of tRNA.

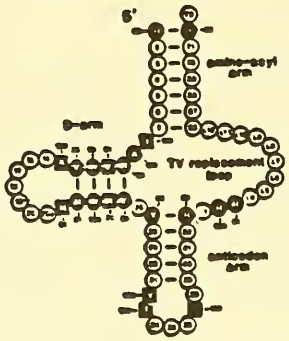


Figure 2. Diagram of the consensus secondary structure of Ascaris suum mitochondrial tRNA-like sequence. The number of nucleotides shown in the D (dihydrouridine) loop and the TV (TψC arm-variable loop)-replacement loop are the maximum numbers observed. Letters within solid squares identify nucleotides or nucleotide combinations that occur with a frequency (accompanying numbers) of greater than 90% among A. suum mitochondrial tRNA-like sequences and are also constant nucleotides in prokaryotic and eukaryotic nuclear-encoded tRNAs. Letters within solid circles identify other nucleotides or nucleotide combinations that occur with a frequency (accompanying numbers) of greater than 80% in A. suum tRNA-like sequences. Nine of these nucleotides (1, 9, 10, 13, 22, 25, 27, 43, and 72) are considered semi-invariant in prokaryotic and eukaryotic nuclear-encoded tRNAs. The numbering system used (1 to 43 and 66 to 73 as given in the open circles) reflects the conventional numbering system used for yeast $tRNA^{phe}$. L1 to 112 denote the maximum of twelve nucleotides observed in the TV-replacement loop of the A. suum mitochondrial tRNA-like sequences. A, adenosine; T, uridine, R, adenine or guanine; Y cytosine or thymine; W, adenine or thymine.

First, if these unusual molecules do fold as indicated in Figure 2, then they appear to contain sequences which appear to be anticodons. The three nucleotides which are situated where anticodons would be expected to be situated (positions 34-36) are unique and compatible with codon usage in the Ascaris mitochondrial protein genes. Second, within the unusual RNA molecules, there is a high frequency of nucleotides or nucleotide combinations that are thought to be either invariant or highly conserved in procaryotic and eucaryotic nuclear-encoded tRNAs. Finally, considerations of the direction of transcription, base composition and relative gene arrangements are all consistent with the hypothesis that these sequences are the tRNAs of Ascaris mitochondria.

A. suum is an obligate parasite and the adult phase is anaerobic. To determine whether the peculiar features of its mitochondrial DNA were related in some way to its mode of life, Dr. Wolstenholme has determined the sequence of most of the mitochondrial DNA of Caenorhabditis elegans, a free-living soil nematode. The results showed that the spacer regions in this DNA are very similar to those of the Ascaris mitochondrial DNA, suggesting that this may be a feature of mitochondrial DNA of all nematodes. Further studies suggest however, that the unusual properties may be limited to nematodes, as the sequence of several putative tRNA regions of the mitochondrial DNA of a flatworm, Fasciola hepatica, appear to conform to the standard clover-leaf structure.

Dr. Wolstenholme is pursuing the analysis of these anomalous coding sequences. He will attempt to determine whether they do code for RNA and if so, whether those RNA molecules have amino acid acceptor activity. Should these unusual structures turn out to be the Ascaris mitochondrial equivalent of tRNAs, they will be extremely interesting to study in terms of the structure-function relationships of tRNA, in that they provide a natural variant set of molecules in which the structure has been drastically altered but the function remains the same (they are, in effect, a genetic engineering experiment which no investigator in his or her right mind would have had the courage to design!). Since tRNAs interact with a number of other proteins and RNAs during protein synthesis, it will be of further interest to detemine the kinds of changes that have occured in other components of the protein synthetic system of Ascaris mitochondria to accomodate the alterations in the tRNAs. Other differences between nematode mitochondria and those of other metazoans, including aspects of gene organization, of the genetic code, and of the mechanism of transcription, may also be of great use in the analysis of basic biological processes. Finally, the existence of such radical differences in the structure of the mitochondrial DNA between one group of organisms and the rest of the metozoa raises first, the evolutionary question of how such major biochemical differences were generated and then fixed and second, the possibility that other unique solutions to fundamental biological issues have occurred in other groups of organisms and will be similarly useful in our further investigation of important biological problems.

Mark S. Guyer, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 1, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Towards a Resolution of the tRNA Identity Crisis" R01 GM16995 (Schulman, L.), Albert Einstein College of Medicine; R01 GM21489 (Murgola, E.), University of Texas System Cancer Center; R01 GM23562 (Schimmel, P.) Massachusetts Institute of Technology; R01 GM37552 (Uhlenbeck, O.), University of Colorado.

To Director, NIGMS

Through: Director, Genetics Program
Chief, Medical and Molecular Genetics Section

The expression of genetic information involves a two-stage process in which the information encoded in the nucleotide sequence of DNA determines the amino acid sequence of a protein. In the first stage, the DNA sequence is transcribed into a polyribonucleotide sequence termed messenger RNA (mRNA). Then, in the second stage, the mRNA nucleotide sequence is used as a guide to assemble amino acid precursors into a specific polypeptide sequence, or protein. A second type of RNA, transfer RNA (tRNA), plays a central role in the process of protein synthesis (translation) by acting as a carrier of the individual amino acids to their correct site along the growing polypeptide chain. There are a large number of tRNA's in every cell. Each can only act as a carrier of one particular amino acid; however, for a number of amino acids there is more than one tRNA which can act as a carrier. In performing out its carrier function, each tRNA must engage in two highly specific recognition processes.

First, the tRNA must be covalently coupled to (charged with) the correct amino acid. tRNA charging is carried out by an enzyme termed an aminoacyl-tRNA synthetase, which forms a covalent bond between the amino acid and the 3' (or acceptor) end of the tRNA molecule (Figure 1). There are twenty such enzymes, one for each of the twenty amino acids which are used for protein synthesis.

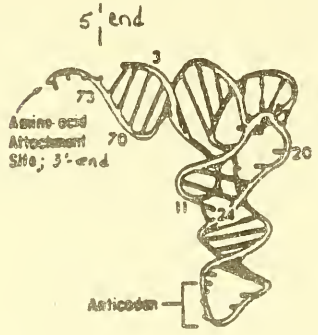


Figure 1. Structure of tRNA. The three-dimensional structure of tRNA, as determined by X-ray crystallography, is shown. The ribose phosphate backbone is depicted as a twisted tube; the base pair interactions are indicated as cross ranges. Because of the folded structure, base pairs can form between nucleotides located at relatively distant positions along the single-stranded RNA. The system for identifying positions takes the 5'-end of the molecule as nucleotide number 1.

Each tRNA molecule must be charged (and therefore recognized) by one synthetase, but not by the nineteen others. Conversely, each synthetase must be capable of recognizing one set of tRNA's (a set may consist of one to six tRNA species), but not nineteen others. Second, the charged tRNA molecule must recognize the specific sequence on the mRNA which corresponds to the site in the growing polypeptide chain at which that particular amino acid is to be added.

The biochemical nature of the second recognition reaction, the mRNA:tRNA interaction, involves base pairing between trinucleotides, one (the codon) in the mRNA and the other (the anticodon, Fig. 1) in the tRNA. This interaction was elucidated in the late 1950's and is a well-understood aspect of the process of gene expression. In contrast, the nature of tRNA recognition by the aminoacyl-tRNA synthetase, although the subject of investigation for many years, has only recently begun to be revealed. A number of NIGMS-supported investigators have made important contributions to the current understanding of the determination of tRNA identity. Their findings suggest that the aminoacyl-tRNA synthetase recognizes specific structural features on the tRNA molecule, that these features are determined by the nucleotide sequence of the tRNA, and that different synthetase-tRNA pairs employ different kinds of recognition features.

The most obvious candidate for a synthetase recognition element in tRNA is the anticodon because this region of the molecule clearly acts as a specificity element in at least one reaction, codon recognition. In some cases, the anticodon does appear to be the primary determinant of tRNA identity (1). For example, work in the laboratory of Dr. LaDonne Schulman (GM16995) several years ago showed that alteration of the CCU anticodon of $tRNA^{Trp}$ to CAU, which recognizes the methionine codon AUG, transformed the tryptophan tRNA into one which is charged with methionine (2). Comparable results were obtained with $tRNA^{Val}$ and, conversely, conversion of the anticodon of $tRNA^{Met}$ to that of $tRNA^{Val}$ changed the charging specificity of the tRNA from methionine to valine. Similarly, several laboratories studying tRNA suppressors (tRNA's in which the anticodon has been changed to recognize the termination codons UAG, UAA, or UGA that normally signal the completion of protein synthesis) showed that the anticodon is also a major recognition element for both the glutaminyl- and trytophanyl-tRNA synthetases.

However, there is clear evidence that the anticodon is not a major determinant of substrate specificity for other aminoacyl-tRNA synthetases. The clearest example is presented by several other suppressor tRNA's, in which the anticodon has been changed from the wild type sequence to one that recognizes a termination codon without affecting the ability of the tRNA to be recognized by the original synthetase. Using techniques such as site-directed mutagenesis or automated gene synthesis, several investigators have recently shown that nucleotides located at sites in the tRNA molecule other than the anticodon are involved in determining the synthetase recognition specificity of some tRNA's. The simplest case is that of a $tRNA^{Ala}$ of Escherichia coli.

Genetic experiments carried out in the laboratory of Dr. Emanuel Murgola (GM21499), and published in 1984 (3), had shown that a change in one base pair in $tRNA^{Lys}$ resulted in mischarging that tRNA with either alanine or glycine instead of lysine. The change involved conversion of the base pair 3:70 (see legend to Figure 1) in the acceptor stem from G:C to G:U. Recently, two laboratories published directed mutagenesis studies that confirmed the critical importance of this base pair in the determination of the identity of the E. coli $tRNA^{Ala}$.

Dr. Paul Schimmel (GM23562) and co-workers used oligonucletide-directed mutagenesis to produce 28 derivatives of a $tRNA^{Ala/CUA}$ (the CUA indicates that this tRNA is a suppressor, the anticodon of which is CUA). A total of 36 nucleotide changes were represented among these 28 derivatives (4). Functional assays indicated that 25 of the 28 were still recognized by the alanyl-tRNA synthetase in spite of the sequence changes. Three variants, however, were no longer recognized by the enzyme. The feature that the three affected variants had in common, and that distinguished them from the other 25 tested, was a change in the G3:U70 base pair in the acceptor stem of the molecule. To study the role of the G3:U70 base pair further, Drs. Schimmel and Hou converted the 3:70 base pair of two other tRNA's, $tRNA^{Cys/CUA}$ and $tRNA^{Phe/CUA}$, to G:U. These two tRNA's were selected, in part, because they are very different from $tRNA^{Ala}$, differing at 38 and 31 positions respectively from $tRNA^{Ala/CUA}$ (out of a total of 76). In spite of all the other differences, however, the conversion of base pair 3:70 to G:U conferred alanine acceptor activity on each of the resulting variants (Figure 2).

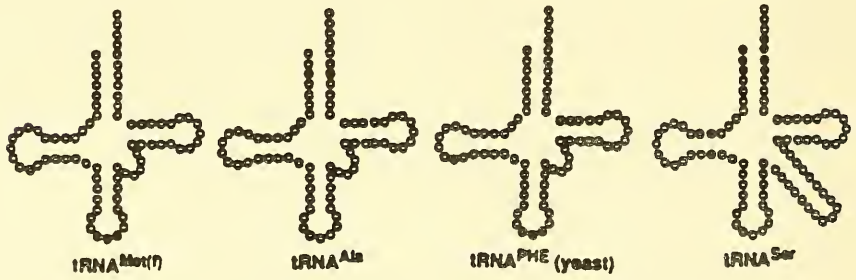


Figure 2. Major identity elements in four tRNA's. The secondary structure of the tRNA's are shown in the cloverleaf configuration; each base is represented as a circle. Filled circles represent nucleotides that have been shown to be required for recognition of the tRNA by the cognate aminoacyl-tRNA synthetase (taken from ref. 1)

A comparable study of $tRNA^{Ala}$ identity was published by Dr. William McClain at about the same time (5). This work also demonstrated that the G3:U70 base pair is a major determinant of the ability of the tRNA to be a substrate for the alanyl-tRNA synthetase (Dr. McClain's work was supported by the NIAID, and the continuation will be supported by NIGMS). In their papers, both groups were careful to point out that their results do not imply that the G3:U70 base pair is the sole determinant of alanine specificity nor that the single base pair 3:70 is the site of synthetase recognition specificity in all tRNA's. Nonetheless, the identification of a simple non-anticodon structural feature of a tRNA molecule as a major element in determining its ability to be specifically recognized by its cognate charging enzyme is a very important result.

Another study of the determinants of tRNA identity currently being carried out uses the $tRNA^{Phe}$ of Saccharomyces cerevisiae. The results of these experiments, from the laboratory of Dr. Olke Uhlenbeck (GM37552), indicate that determination of $tRNA^{Phe}$ identity is different from any of the cases described above. As in Dr. Schimmel's work, this project began with an extensive analysis of a large number of mutant tRNA's, in this case made primarily by automated gene synthesis. Within the collection of mutants synthesized, Dr. Uhlenbeck and his co-workers changed the nucleotides at all but five of the positions that are not conserved among all yeast tRNA's (there are 17 conserved nucleotides among yeast tRNA's; the primary function of these appears to be structural, the maintenance of the folded conformation of the tRNA molecule). On the basis of the behavior of the mutant tRNA's, these investigators identified five bases that appeared to confer specificity for the phenylalanyl-tRNA synthetase. These were the three anticodon bases G34, A35, and A36, plus non-anticodon bases, G20 and A73 (Fig. 2). Any change at any of these positions was found to reduce the rate of charging by the phenylalanyl-tRNA synthetase by about ten-fold; these effects were multiplicative. Furthermore, introducing the five bases into other yeast tRNA's (e.g., $tRNA^{Met}$, $tRNA^{Arg}$, and $tRNA^{Tyr}$) allowed the new variants to be recognized by the phenylalanyl-tRNA synthetase at efficiencies of 65% to greater than 100% of that of $tRNA^{Phe}$. The involvement of several nucleotides also appears to be the case for $tRNA^{Ser}$, although the anticodon is probably not involved in this instance (Normanly and Abelson, cited in ref. 1; Fig. 2).

From these and other studies, several generalizations have emerged (1).

(a) Only a small number of nucleotides are required to establish the identity of a tRNA.

(b) The anticodon may or may not be a specificity element.

(c) Non-productive interactions with non-cognate synthetases must be prevented by "negative" specificity features.

It is also interesting to note that the some specificity elements have been conserved in evolution (the G3:U70 base pair has been found in every cytoplasmic $tRNA^{Ala}$ that has been sequenced, from archaebacteria to eucaryotes), while others have not (position 20 is G in yeast $tRNA^{Phe}$, but U in E. coli $tRNA^{Phe}$).

Although the term "second genetic code" has been suggested to describe the tRNA-synthetase recognition system, this is not appropriate. In contrast to the "genetic code," which refers to the codon:anticodon recognition mechanism, the results obtained to date indicate there is not a strict one to one correspondence between a tRNA nucleotide sequence and a decoding site on the synthetase, as the term "code" implies.

The progress in the analysis of the decades-old question of what determines tRNA identity has, in large part, been made possible by recent technological advances, primarily oligonucleotide synthesis. The fact that major advances have come simultaneously from several laboratories suggests that this is a scientific problem whose time has come.

Mark S. Guyer, Ph.D.

1. Schulman, L. H. and J. Abelson. (1988) Science. 240: 1591.

2. Schulman, L. H. and H. Pelka. (1983) Proc. Nat. Acad. Sci., U.S.A. 80: 6755.

3. Prather, N.E., et al. (1984) J. Mol. Biol. 172: 177.

4. Hou, Y.-M. and P. Schimmel. (1988) Nature. 333: 140.

5. McClain, W. and K. Foss (1988) Science. 240: 793.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date October 6, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT - "Pseudoknots: A New Twist in RNA-Protein Interactions"; R01 GM29048 (Draper, D.) The Johns Hopkins University

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 10-13-89
Chief, Molecular & Medical Genetics Section, Genetics Program, NIGMS 10-7-89

The formation of specific RNA-protein complexes is essential for many aspects of gene expression. The ribosome, the site of protein synthesis, is an example of an RNA-protein complex that is relatively stable. Other, shorter-lived, RNA-protein interactions are responsible for events such as regulation of messenger RNA translation and aminoacylation of transfer RNA. The importance of these interactions has led to numerous attempts to define protein recognition sites in RNA. These studies, however, have proved to be much more challenging than analogous studies of DNA-protein interactions. In the case of double-stranded DNA, proteins usually bind to short contiguous sequences that can be identified by a number of standard techniques. In contrast, recognition sites in RNA are often created by three dimensional folding of the polynucleotide chain. If the nucleotides that contact the protein are highly dispersed within the primary sequence, as is often the case, identification of these sites in a large RNA can be extremely difficult.

NIGMS grantee Dr. David Draper at The Johns Hopkins University has been employing a novel approach to investigate the interactions between ribosomal proteins and RNA. One method of identifying recognition sites in RNA is to compare the sequences of two or more RNAs that bind to the same protein. Ribosomal RNAs from evolutionarily related organisms have been used in this manner. Although consensus sequences can be derived from such comparisons, it is often difficult to distinguish functional similarities from those due to protein recognition. Dr. Draper has chosen to investigate those ribosomal proteins that, because they also bind messenger RNA, interact with two functionally dissimilar RNAs. His studies with the ribosomal protein S4 from E. coli have been particularly illuminating as they have revealed a new and unusual topological feature of RNA recognition sites.

The genes for the 52 ribosomal proteins of E. coli are grouped into 19 operons that are distributed throughout the genome. The coordinated expression of these genes is achieved by an autogenous mechanism that operates at the level of translation. In most of these operons, one of the encoded ribosomal proteins also serves as a translational repressor. In the alpha operon (Fig. 1), ribosomal protein S4 performs this function. The S4 protein normally binds to 16S rRNA as one of the initial steps in assembly of the 30S ribosomal subunit. However, if the rate of expression of the alpha operon outstrips synthesis of the other ribosomal components, free S4 protein binds to the 5'-leader region of the alpha operon messenger RNA and represses further translation. For reasons that are not yet understood, the one non-ribosomal gene in the alpha operon is not regulated by S4.

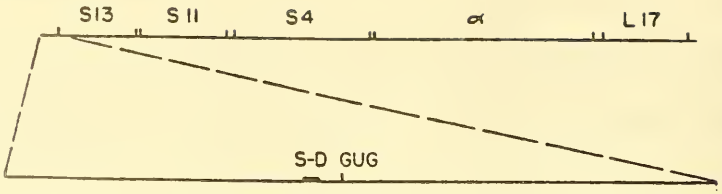


Figure 1. Map of the alpha operon of E. coli. The top line indicates the positions of the five genes of the operon. S13, S11, and S4 are components of the 30S ribosomal subunit. alpha encodes a subunit of RNA polymerase. L17 is a component of the 50S ribosomal subunit. The bottom line represents an expanded view of the 5'-leader region. S-D and GUG indicate the positions of the Shine-Dalgarno sequence and the initiation codon of the S13 gene, respectively. The initiation codon begins at nucleotide 95, numbered from the 5'-terminus of the messenger RNA.

Dr. Draper has been exploring the messenger and ribosomal RNA recognition sites of the S4 protein. Studies with fragments of alpha operon mRNA indicated that S4 binds to a 130 nucleotide segment that is located close to the 5'-end of the molecule. This segment contains an untranslated 5'-leader region where sites that regulate translation usually occur. The secondary structure of the 5'-end of the alpha operon mRNA was initially probed with endoribonucleases that are specific for double- or single-stranded RNA. Based on the locations of the helical and single-stranded regions, Dr. Draper deduced a preliminary model for the 5'-end of the mRNA in which the polynucleotide chain is folded into a compact structure known as a pseudoknot. As illustrated in Fig. 2, pseudoknots occur when an RNA chain contains a hairpin, or stem and loop structure, in which bases in the loop are paired with complementary sequences in a nearby part of the molecule. In the absence of steric constraints, the two paired regions can form a quasi-continuous helical segment that can be stabilized by base stacking interactions. Pseudoknots were first recognized at the 3'-termini of certain plant viral RNAs and more recently several have been tentatively identified in other viral and bacterial RNAs.

The proposed pseudoknot structure for alpha operon mRNA was tested by measuring the binding of S4 protein to variants created by site-directed mutagenesis. Dr. Draper reasoned that if two segments of the RNA formed a helical region that was essential for binding S4, mutations that disrupt base pairing between these segments should weaken binding. By the same logic, binding should be restored by introduction of compensating second site mutations that restore base pairing. The results of these analyses confirmed the pseudoknot structure and revealed the existence of a third helical segment that had not been part of the original model. The resulting structure, which Dr. Draper has termed a "double pseudoknot" (See Fig. 3), is the most complex of the pseudoknot structures that have been detected to date.

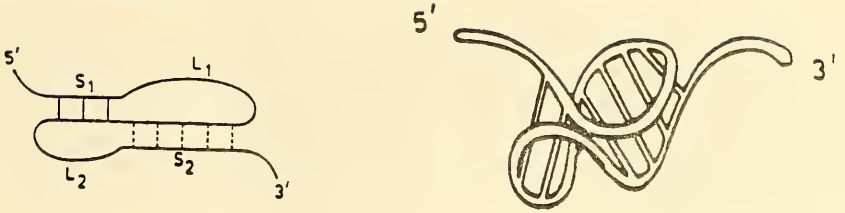


Figure 2. Two representations of an RNA pseudoknot. In the schematic diagram on the left, S1 and S2 indicate helical stem regions formed by Watson-Crick base pairing and L1 and L2 indicate single-stranded loops that connect the helical regions. The three-dimensional folding on the right illustrates the formation of a quasi-continuous helix involving S1 and S2.

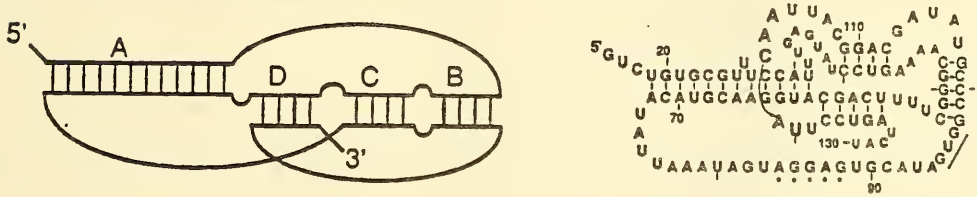


Figure 3. A schematic view of the double pseudoknot structure deduced for alpha operon mRNA leader region (left) and the primary and secondary structure of alpha mRNA in the double pseudoknot configuration (right). The letters in the schematic diagram on the left represent the 4 helical regions shown on the right. A designates bases 19-31/61-72; B, 40-43/108-111; C, 48-52/97-102; D, 56-59/123-126. Helical regions B and C are topologically equivalent. The curved arrow indicates the normal phosphodiester bond between bases 118 and 119. The Shine-Delgarno sequence is denoted by asterisks; the initiation codon is underlined.

Dr. Draper has found that the S4 binding site in 16S ribosomal RNA has been somewhat less amenable to analysis. Studies with fragments of 16S rRNA have indicated that the region between nucleotides 38 to 500, which comprises most of the 5'-third of the molecule, is required for S4 binding. The secondary structure of this region, which contains a number of short helical regions (See Fig. 4), has been extensively studied by MIGMS grantee Dr. Harry Noller (GM 17129). Surprisingly, a comparison of the S4 binding regions of 16S rRNA and alpha mRNA revealed little homology at the level of either primary or secondary structure. Moreover, Dr. Draper was able to demonstrate that the only readily apparent structural similarity could be altered in either molecule without affecting S4 binding. In spite of this lack of similarity, alpha mRNA and 16S

rRNA compete for S4 binding, and the binding activities of both RNAs exhibit an unusual enhancement in the presence of chloride ion. It is thus likely that S4 recognizes both molecules by the same mechanism and that both molecules can adopt tertiary structures that create identical binding sites. Due to its large size and complex structure, 16S rRNA has the potential to create binding sites that involve a number of different helical segments. The formation of a compact double pseudoknot structure may offer a means by which the smaller alpha mRNA leader region can create an equivalent combination of RNA strands.

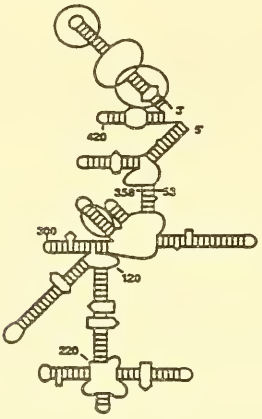


Figure 4. Primary and secondary structure of the 5'-domain of E. coli 16S rRNA.

The exact mechanism by which S4 protein represses translation of the alpha operon is not known. Translational repressors generally act either by preventing binding of ribosomes to mRNA or by interfering with the subsequent movement of ribosomes along the message. As shown in Fig. 3, the Shine-Dalgarno sequence, which forms part of the ribosome attachment site, and the first 7 or 8 codons of the S13 gene both occur within the region that forms the double pseudoknot. Thus, by binding to and stabilizing the pseudoknot structure, S4 protein could inhibit translation by either mechanism. Dr. Draper is continuing to explore the structural features and functional significance of the double pseudoknot. He has recently found that up to four bases can be deleted from the longest single-stranded loop (bases 73 to 97 in Fig. 3) without affecting S4 binding. One such mutant, however, no longer exhibited translation repression, suggesting that S4 binding alone is not sufficient to inhibit translation. Dr. Draper suspects that S4 may induce subtle structural changes in the pseudoknot region and that regulation depends on the conformation adopted by the mRNA, rather than binding of S4 *per se*.

Pseudoknots do not appear to be a common feature of RNA recognition sites. Dr. Draper has investigated several other ribosomal proteins that, like S4, serve as translational repressors. In each of these cases, the ribosomal and messenger RNA recognition sites exhibit considerable similarity. These sites, which consist of a short helical segment with a characteristic single-stranded loop,

are much smaller than the S4 recognition sequence and there is no evidence for the kind of intramolecular base pairing that occurs in a pseudoknot. Relatively little information is available on the functional significance of the handful of other pseudoknots that have been identified. Recently, however, a pseudoknot was discovered to serve as a signal for ribosomal frameshifting in an avian coronavirus. In this instance, the pseudoknot induces a -1 shift in the reading frame during translation, resulting in the production of a single protein from two partially overlapping genes. Thus, although they may occur infrequently, it seems likely that pseudoknots will be found to serve a variety of functional roles in RNA metabolism.

Marc Rhoades, Ph.D.

References

1. Draper, D.E. 1989. How do proteins recognize specific RNA sites? Trends Biochem. Sci. 14:335-338.

2. Tang, C. and D. Draper. 1989. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. Cell. 57:531-536.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 30, 1988

From Program Administrator, Genetics Program, NIGMS

Subject Research Highlight: "A Phage T4 Gene Interruption That Is Not Removed Before Translation: A 'Giant Leap' for the Ribosome"; Wai Mun Huang, University of Utah Medical Center, Salt Lake City, UT; RO1 GM-21960,

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 12-2-88

It has been known for some time that coding sequences of genes of higher eukaryotes are generally interrupted by untranslated sequences called introns. These introns are present in the initial RNA transcript of the gene, but are subsequently removed before the messenger RNA is translated by ribosomes. The removal of introns is usually accomplished by particles composed of both RNA and protein (called snRNPs, for small nuclear ribonucleoproteins) that are found in the nucleus. The excision reaction involves specific cleavage of the RNA at one end of the intron, followed by splicing. In some cases, however, introns are self-splicing, and the cleavage and rejoining reactions are catalyzed by the RNA itself.

It is still not clear whether introns reflect some essential aspect of genome structure or RNA processing in higher eukaryotes, whether they provide advantages in facilitating the evolution of new protein structures, or whether they are simply remnants of early stages of evolution that are not worth the effort to remove from genes. For some time it was believed that introns were present only in eukaryotic genes, and not in those of bacteria, and very many bacterial genes were sequenced and shown not to be interrupted. In fact it was predicted that introns would not be found in bacteria because, at least in Escherichia coli, ribosomes begin translating mRNA as soon as its 5'- end has been synthesized, and this allows no time for splicing. In eukaryotes transcription and translation are separated by the nuclear membrane, so the entire RNA can be synthesized and spliced in the nucleus before being transported to the cytoplasm for translation.

Surprisingly, introns were discovered several years ago in a few genes that function in bacteria, though all are genes of bacteriophages. Self-splicing introns have been demonstrated in three genes of E.coli bacteriophage T4, one encoding thymidylate synthase, one encoding a subunit of ribonucleotide reductase, and one of unknown function. A self-splicing intron has also been discovered in the Bacillus subtilis phage SPO1. All of these sequences are removed from transcribed RNA by similar self-splicing mechanisms requiring no additional enzymes.

Dr. Wai Mun Huang and coworkers have now discovered, in another bacteriophage T4 gene, a unique untranslated sequence, one which is not excised from the message before it is translated. The gene, gene 60, encodes one of three subunits of the enzyme DNA topoisomerase, which relaxes supercoiled DNA and is required for DNA replication. Dr. Huang's group cloned the gene in an expression vector in order to obtain large amounts of the protein product, p60, for biochemical studies. When they determined the DNA sequence of the cloned gene, they were surprised to find a stop codon after the 46th codon, which would make a protein much too short. Yet the cloned gene was found to direct the synthesis of normal p60, which could assemble into active topoisomerase. Downstream in the sequence, some 50 nucleotides away, was the rest of the sequence required to encode the entire p60.

Suspecting the presence of an intron, Dr. Huang used reverse transcriptase with the appropriate deoxyoligonucleotide primer to determine the sequence of gene 60 messenger RNA from infected cells. The sequence was the same as that of the cloned gene, still containing the interruption. By this method, any spliced mRNA would have been detected as an alternative sequence beyond the splice site. No such sequence was detected, though as little as 5% splicing would have been apparent. Incubating the RNA in vitro under conditions that allow self-splicing of other T4 introns produced no processing of the gene 60 message. In order to determine whether translation might be limited to a small fraction of the message population that is spliced, Dr. Huang inserted gene 60 sequences, including the interruption, into the gene for beta-galactosidase, in the correct reading frame so that translation of beta-galactosidase would require prior translation of the interrupted gene 60 sequence. This did not decrease the yield of beta-galactosidase, as would have been the case if splicing were required and limited to 5% of the message population.

These experiments and the fact that the interruption in gene 60 lacks consensus splice site sequences present in known introns led Dr. Huang to conclude that splicing of the mRNA is unlikely (but not rigorously excluded). In order to produce p60 protein, therefore, the ribosome must apparently bypass the interruption, which is 50 nucleotides in length. Translation of other genes in E.coli has been found to require ribosome frameshifting, and frameshifts of up to 6 nucleotides have been demonstrated, but never before has a "leap" as long as 50 nucleotides been observed in any system. In order to explain this ribosome bypass, Dr. Huang has proposed that the RNA assumes a secondary structure that brings the codons for the adjacent amino acids into juxtaposition, allowing the ribosome to skip across the interruption (see Fig. 1).

Dr. Huang suggests that this interruption presents an additional potential site for regulation of topoisomerase synthesis. However, it is interesting to note that the interruption is not found in the homologous gene of the closely related bacteriophage T2. In T2, sequences equivalent to T4 genes 39 (which encodes another subunit of topoisomerase) and 60 are joined, and the combined message contains one continuous open reading frame without interruption. This is true even though the protein sequence at the site of interruption in T4 (Asp-Gly-Leu) is conserved in T2.

Whatever physiological significance ribosome bypass may have in this system, now that it has been recognized that ribosomes can bypass mRNA sequences during translation, other examples will probably be discovered before long. This unique observation provides another example of the principle that gene expression need not be a simple read-out of the genetic information. This information can be modified by the expression machinery, so that not only the level and timing of expression, but also (in principle) the structure of the protein produced can be altered in response to physiological conditions.

Stephen R. Fahnestock

Reference:

Huang, W.M., Ao, S.-Z., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D., and Fang, M. (1988) Science 239, 1005-1012.

(Figure 1)

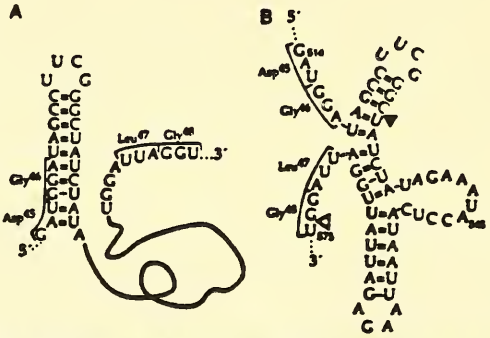


Fig. 8. Proposed secondary structures of T4 gene *60* mRNA in the interrupted region. The filled and unfilled triangles mark the positions of the major and minor AMV reverse transcriptase stops, respectively. Nucleotide numbers (indicated at the lower right-hand corners of a nucleotide) are defined according to Fig. 3A. An 8-base region, starting from position 545, is complementary to the gene *60* translation initiation region (Fig. 3A, position 369 to 376). The significance of this potential pairing remains to be determined.

C. THE GENETICS OF DEVELOPMENT

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date January 11, 1988

From Deputy Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "A Mother's Work Is Never Done: Role of Localized Maternal Messenger RNAs in Embryonic Development," RO1 GM 32921-04 (Melton, D.A., Harvard University)

To Director, NIGMS
Through: Director, Genetics Program, NIGMS

Embryologists have long puzzled over the mechanisms that cause the daughter cells of a single fertilized egg to assume distinct developmental fates. In 1925, E.B. Wilson proposed that a differential distribution of maternal factors in the egg might explain why cells arising from different parts of the egg develop into different embryonic structures. Wilson's suggestion is supported by much indirect evidence for the uneven distribution of maternal factors in the egg's cytoplasm, and it has been proposed that proteins and/or messenger RNAs (mRNAs) could be the factors that determine cell fate.

In an attempt to discover the role of maternal factors in development, Dr. Douglas Melton and his colleagues set out several years ago to determine whether localized maternal mRNAs in fact exist in eggs (1). For their studies they chose Xenopus eggs which exhibit obvious differences in morphological, biochemical, and developmental properties along the animal-vegetal axis. The animal (top) hemisphere of the egg is destined to form ectoderm, and the vegetal (bottom) hemisphere is destined to form endoderm. Dr. Melton's group first prepared from unfertilized eggs a cDNA library which would be expected to contain DNA complementary to most of the egg's maternally derived mRNAs. The cDNAs were radioactively labeled and allowed to hybridize to RNA that had been isolated from either the animal or vegetal regions of unfertilized eggs. In most cases, RNAs that hybridized to the labeled cDNA were present equally in both the animal and vegetal regions as would be expected if they were uniformly distributed throughout the egg. However, a small percentage of the maternal mRNAs appeared to be localized in one hemisphere or the other. Three cDNA clones for animal-specific (An1, An2, An3) mRNAs and one for vegetal-specific (Vg1) mRNA were analyzed in more detail. They found that these localized mRNAs do not become redistributed by the cytoplasmic movements that are known to follow fertilization in Xenopus and they maintain their positions during cell division. These results demonstrate that specific daughter cells receive certain maternal mRNAs as a consequence of the initial cytoplasmic distribution of these mRNAs in the egg. Interestingly, An1 and Vg1 are strictly maternal (no transcripts are made by the embryo) whereas An2 and An3 are also transcribed throughout embryogenesis.

Having shown that some maternal mRNAs are localized in eggs, Dr. Melton was interested in learning what, if any, developmental function these mRNAs or their protein products have. So far, he and members of his laboratory have

identified the products of two of the mRNAs. Their findings suggest that localized maternal mRNAs can play very different roles in development.Dr. Melton determined the DNA sequence corresponding to Vg1 mRNA and compared its deduced amino acid sequence to those of known proteins (2). Surprisingly, Vg1 mRNA encodes a protein homologous to members of the transforming growth factor-β (TGF-β) family. Dr. Melton had previously shown (3) that Vg1 mRNA is initially localized in a thin crescent at the vegetal pole of mature oocytes, but that later, in the unfertilized egg and the early embryo, Vg1 mRNA spreads toward the equator to form a broader band filling most of the vegetal region. The strict localization to the vegetal region implies that the function of Vg1 mRNA is unique to cells in that region, and the finding that the Vg1 protein is a growth factor suggests that it acts as an intercellular signal. Other members of the TGF-β family have been shown to stimulate or inhibit cell proliferation, affect development, and potentiate the effects of other growth factors.

It is well known that the mesodermal tissue layer arises from ectoderm (animal region) only if it is induced by endoderm (vegetal region) or extracts of endodermal tissue. Dr. Melton has therefore hypothesized that Vg1 protein, secreted by vegetal cells, interacts with cells in the more animal region of the early embryo to induce them to form mesoderm. In this way, the localized Vg1 mRNA could play a critical role in specifying the structure of the early embryo.

Using a complementary approach, Dr. Marc Kirschner (4) (P01 GM 31286) has recently shown that bovine fibroblast growth factor (FGF) and a closely related protein from Xenopus have a limited capacity to induce the expression of muscle actin (muscle is derived from mesoderm) in cells from the animal region of Xenopus embryos. The presence of TGF-β significantly raises the level of expression of actin induced by FGF. This provides further support for the role of a TGF-β type protein in the induction of mesoderm.

The second localized maternal mRNA sequenced by Dr. Melton's laboratory, An2, was found to encode a protein with a strong homology to the alpha subunit of mitochondrial ATPase (5). This protein is encoded by a gene in the nucleus and is transported from its site of synthesis in the cytoplasm to mitochondria. The discovery that a localized mRNA encodes a subunit of ATPase was surprising, since one would not expect such a molecule to play a regulatory role in development. However, the animal region is known to have a higher rate of respiration corresponding to its higher rate of cell division and cell motility during gastrulation. It is therefore likely that the localization of An2 mRNA could exert an effect on cell fate by affecting respiration, ion flow, or metabolism.

Dr. Melton's studies, like all good science, have followed a logical path but have uncovered in the process some surprises which open up new areas of inquiry. Clearly, because of their apparently diverse roles, it will be important to investigate the functions of other localized maternal mRNAs, and it will be of great interest to characterize the precise roles of the Vg1 and An2 protein products, as well as the times at which they are translated. It

has been recognized for a long time that the egg provides more than just half of the offspring's genetic complement. Dr. Melton's studies are beginning to pinpoint the mother's other contributions to the development the embryo.

Judith H. Greenberg, Ph.D.

References:

1. Rebagliati, M.R., Weeks, D.L., Harvey, R.P., and Melton, D.A. Identification and cloning of localized maternal RNAs from Xenopus eggs. Cell 42: 769-777, 1985.

2. Weeks, D.L. and Melton, D.A. A maternal mRNA localized to the vegetal hemisphere in Xenopus eggs codes for a growth factor related to TGF-β. Cell 51: 861-867, 1987.

3. Melton, D.A. Translocation of a localized maternal mRNA to the vegetal pole of Xenopus oocytes. Nature 328: 80-82, 1987.

4. Kimelman, D. and Kirschner, M. Synergistic induction of mesoderm by FGF and TGF-β and the identification of an mRNA coding for FGF in the early Xenopus embryo. Cell 51: 869-877, 1987.

5. Weeks, D. L. and Melton, D.A. A maternal mRNA localized to the animal pole of Xenopus eggs encodes a subunit of mitochondrial ATPase. Proc. Nat'l. Acad. Sci. USA 84: 2798-2802, 1987.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 7, 1989

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT - "Sex Determination: Viva la Differential mRNA Splicing" R01 GM37731 (Baker, B.) Stanford University, R01 GM36579 (Belote, J.) Syracuse University, R01 GM 23468 (Cline, T.) Princeton University, R01 GM36549 (McKeown, M.) Salk Institute, R01 GM 25976 (Schedl, P.) Princeton University

To Director, NIGMS

In Drosophila, the primary determinant for whether a somatic cell will be phenotypically male or female is the ratio between the number of X chromosomes and sets of autosomes. A cell with two X chromosomes and two sets of autosomes (2X:2A) expresses female characteristics, and a cell with one X chromosome and two sets of autosomes (1X:2A) expresses male characteristics. However, the X:A ratio acts through a regulatory hierarchy of genes, which ultimately specify whether a cell expresses a male or female phenotype. During the last decade, genetic studies have demonstrated that the X:A ratio acts through a gene called sex-lethal, which is the first gene in the regulatory pathway shown below:

X:A ratio → sex-lethal (sxl) → transformer (tra), transformer-2 (tra-2) → doublesex (dsx) → terminal differentiation

The normal function of the sxl, tra, and tra-2 genes is to cause dsx to be expressed in a manner that leads to female differentiation. If any of the genes in the pathway leading to dsx do not function normally, dsx is expressed in a manner that leads to male differentiation. Expression of the male phenotype is therefore the "default mode."

The first clue to how this regulatory hierarchy works was discovered in 1987. At that time, two new NIGMS grantees, John Belote and Michael McKeown, who had both been postdoctoral fellows in the laboratory of another NIGMS grantee, Dr. Bruce Baker, found that the tra gene, which they cloned, encodes a messenger RNA (mRNA) that is spliced in one way in males and in a different way in females (1). Males produce an mRNA transcript of 1.1 kb in length while females produce both a 1.1 kb transcript and a 0.9 kb transcript. The size difference is due to alternative splicing, leading to the removal of different introns from the mRNA. As a result, the female-specific 0.9 kb transcript contains a large open reading frame that encodes a polypeptide of 197 amino acids. In contrast, the 1.1 kb transcript in males has no large open reading frame.

Next, Drs. Baker, Belote, and McKeown tested the hypothesis that the products of genes in the sex determination pathway regulate genes acting downstream (later in the regulatory pathway) by influencing the splicing pattern of their mRNA (2). To do this, they examined the splicing pattern of wild type tra mRNA in XX flies with mutations in either the sxl, tra-2, or dsx locus. They found that

chromosomally female (XX) flies with a mutation in the sxl gene show a male-specific pattern of tra transcripts. In contrast, in chromosomally male (XY) flies, a mutation in sxl does not change the pattern of splicing from the normal male pattern, and the amount of tra mRNA transcribed remains the same as in males with a normal sxl gene. This showed that the proper expression of sxl affects only the female-specific splicing of the tra mRNA and not the male-specific splicing or the overall level of tra mRNA synthesized. Their results also demonstrated that the sxl gene must act upstream from the tra gene (earlier in the regulatory pathway), as suggested by genetic studies. This group performed similar experiments to test the effect of mutations in the tra-2 and dsx genes on splicing of tra mRNA and of mutations in the sxl, tra, and tra-2 genes on splicing of dsx mRNA. In all cases, the molecular experiments confirmed the order of the genes in the regulatory hierarchy that was originally suggested by genetic experiments and showed that feedback regulation of the tra and dsx genes from downstream genes does not occur. Most importantly for this story, they demonstrated that control of the sex determination pathway is accomplished by differential mRNA splicing rather that by regulation of mRNA transcription.

Another NIGMS grantee, Dr. Thomas Cline, has been interested in the function of the sxl gene, the first gene in the pathway. Several years ago, he showed, in genetic experiments, that sxl not only affects the genes acting downstream from it, but also displays autoregulatory activity, which provides a positive feedback loop that stably maintains sxl expression in the female mode. Dr. Cline and another NIGMS grantee, Dr. Paul Schedl, have recently analyzed the structures of the sxl mRNA transcripts in males and females (3). As in the case of tra, the two sex-specific transcripts are alternatively spliced, leading in the female to a transcript capable of encoding a 354 amino acid protein. In males, the transcript contains an extra exon, which introduces a stop codon, and the open reading frame is interrupted so that only a short polypeptide, if any, is produced. Drs. Cline and Schedl found that the protein that is predicted to be encoded by the female-specific sxl transcript has a sequence similar to that of RNA binding proteins. Proteins in this family include several heterogeneous nuclear RNP (hnRNP) components that associate with mRNA precursors in the nucleus, an RNP component that binds to poly(A) RNA, and a component of U1 small nuclear RNP (snRNP) that has been implicated in splicing. This suggested that the sxl protein might also be an RNA binding protein. Thus, in females, the sxl protein might activate tra by directing the splicing of tra mRNA into the functional mode. In males, the nonfunctional mode might occur by default in the absence of the sxl product. Similarly, the sxl protein could be responsible for its autoregulatory activity in maintaining sxl function by acting directly on its own mRNA.

Dr. Baker had shown previously that dsx, like tra and sxl, produces an mRNA that is alternatively spliced in males and females. Unlike the transcripts from those genes, however, both male- and female-specific dsx transcripts contain open reading frames capable of encoding proteins (4). His laboratory subsequently demonstrated that the products of both dsx transcripts also resemble RNA binding proteins (5). The two transcripts share three common exons at the 5' end, but the female transcript contains one unique 3' exon and the male transcript two unique 3' exons. It is well known that when an intron is removed from an mRNA

transcript, the nucleotide sequences at the ends of the two exons surrounding the intron are brought together in a particular way. The end of one exon is referred to as the splice donor site, and the end of the other exon is referred to as the splice acceptor site. When Dr. Baker examined the splice donor and acceptor sites of the common and unique exons of the dsx transcripts, he found that most of these sites resemble those found in other Drosophila genes. The female-specific acceptor site, however, is different from most other acceptor sites, and by this criterion would be expected to be a very poor acceptor. Since use of the female-specific acceptor depends on the activity of tra and tra-2, Dr. Baker believes that the products of these genes might act to redirect the normal preference of the splicing machinery of the cell from the stronger "consensus" male acceptor sequence to the weaker female acceptor sequence.

Although highly suggestive, Dr. Baker's work provides only circumstantial evidence that an upstream sex determination gene influences the splice acceptor sites used by the downstream gene. However, a very recent paper by Drs. Belote and McKeown provides proof that, at least in the case of one sex determination gene, the upstream gene product actively influences a female-specific choice of splice acceptor site (6). These investigators suggested that the alternative splicing observed in the tra mRNA represents a case of acceptor splice site competition. In the absence of sxl activity (as occurs in XY flies or in XX flies with a defective sxl gene), the non-sex-specific splice site of tra mRNA is always used. In the presence of sxl activity (as normally occurs in XX flies), the competition between the two splice acceptor sites is shifted such that the female-specific site is used part of the time.

They further suggested that the shift in competitive balance could occur by two different models. Either the sxl protein activates the otherwise inactive female-specific splice acceptor site, or else the sxl protein blocks the use of the non-sex-specific splice acceptor site. To test the alternative models, they first determined that the information necessary for sex-specific acceptor splice site utilization is contained within the tra intron. This discovery enabled them to modify that intron in various ways and assess whether sex-specific splicing would still occur. The blockage model predicts that removal of the non-sex-specific acceptor site should have the same effect as blocking that site with the sxl protein; both should lead to the use of the female-specific splice site. In contrast, the activation model predicts that the female-specific splice site would not be used in the absence of the sxl protein, so removal of the non-sex-specific acceptor site should have no effect on splicing of the tra intron. In fact, they found that XY flies in which the non-sex-specific acceptor splice site of the tra intron was deleted show substantial female differentiation. Other experiments confirmed the validity of the blockage model. For example, the investigators changed three of the twelve nucleotides in the non-sex-specific acceptor site where the sxl protein would be expected to bind. If the blockage model is correct, altering this region should prevent binding of the sxl protein, the tra intron should not be responsive to the sxl protein, and female flies containing this intron should not express a female phenotype. In fact, such XX flies differentiate with a male phenotype. Also of importance in confirming the blockage model is the fact that the putative binding site for the sxl protein

is the same in the sxl mRNA and the tra mRNA, both of which are thought to be regulated by the same protein.

Taken together, the results of these studies have produced an extraordinarily complete picture of how sex determination occurs, at least in flies. Undoubtedly, it will not be long before some of the missing details, such as the precise binding sites of the proteins that regulate differential splicing, are worked out. What is clear even now, however, is that the major difference between male and female flies lies in the control of splicing. Although alternative splicing has been documented in many systems besides sex determination, it is highly unusual that splicing, rather than transcriptional regulation, plays the primary role in any developmental pathway. One explanation that has been advanced is that, since the regulatory network for sex determination is so essential and is required continuously throughout the life of the fly, it may be advantageous to utilize a separate control mechanism, distinct from the transcriptional controls that appear to regulate so much of the the development of the Drosophila body pattern. Whatever the explanation, viva la differential splicing.

Judith H. Greenberg

Judith H. Greenberg, Ph.D.

References

1. Boggs, R.T., Gregor, P., Idriss, S., Belote, J.M., and McKeown, M. (1987) Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell 50: 739-747.

2. Nagoshi, R.N., McKeown, M., Burtis, K.C., Belote, J.M., and Baker, B.S. (1988) The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster. Cell 53:229-236.

3. Bell, L.R., Maine, E.M., Schedl, P., and Cline, T.W. (1988) Sex-lethal, a Drosophila sex determination gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. Cell 55:1037-1046.

4. Baker, B.S. and Wolfner, M.F. (1988) A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in Drosophila melanogaster. Genes Dev. 2:477-489.

5. Burtis, K.C. and Baker, B.S. (1989) Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell 56:997-1010.

6. Sosnowski, B.A., Belote, J.M., and McKeown, M. (1989) Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. Cell 58:449-459.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date June 28, 1988

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Of Frogs and Flies: Evidence that Homeobox-Containing Proteins in Xenopus Affect Pattern Formation", RO1 GM 32921 (Melton, D.), Harvard University

To Director, NIGMS

An early step in the formation of the body pattern in Drosophila is the division of the embryo into segments. Over the last several years, a number of genes which direct the morphogenetic events leading to segmentation in Drosophila have been identified and cloned, and their functional roles have been characterized. It was quickly realized that genes which regulate segmentation and other aspects of pattern formation in Drosophila share a common sequence of 180 nucleotides, which was named the homeobox, and which encodes the DNA-binding region of these regulatory proteins. Presumably, proteins encoded by these genes interact with other genes in a complex hierarchical manner to specify the body pattern of the organism.

Shortly after their discovery in Drosophila, homeobox-containing genes were detected in other species, including vertebrates. This led to widespread speculation about their roles in organisms that are not so obviously segmented as flies. In mice, for example, it was shown that homeobox-containing genes are expressed in a tissue-specific and stage-specific way. However, because no vertebrate animals mutant for any homeobox-containing gene have been identified, nothing has yet been learned about whether these genes might function in vertebrate development in a similar way to what was suggested in flies.

Recently, Douglas Melton and a postdoctoral fellow in his laboratory, R.P. Harvey, reported results that provide evidence that at least one homeobox-containing gene in the frog Xenopus is important in the development of segmented structures (1).

In 1986, Dr. Melton's laboratory isolated from Xenopus a gene that contains a homeobox sequence closely related to that of the antennapedia gene of Drosophila (2). This gene, called Xhox-1A, is expressed beginning at the mid-gastrula stage. In systems like Xenopus, where there are no known mutations of homeobox-containing genes, one strategy for studying the function of a protein is to deregulate its expression by supplying embryos with exogenous copies of a cloned gene. This approach has been successfully used to deregulate genes in both Drosophila and mice. However, there is currently no method for introducing genes into the chromosomes of frogs, so Drs. Harvey and Melton decided instead to microinject frog eggs with messenger RNA (mRNA), which they synthesized in vitro from the Xhox-1A gene. As one control, they synthesized mRNA from the Xhox-1A gene in which a deletion was made such that the resultant protein would lack a complete homeobox and would contain a different amino acid sequence directly adjacent to the C-terminal end of the homeobox. As a second control, they synthesized Xenopus beta-globin mRNA.

They then microinjected these mRNAs into Xenopus eggs. They demonstrated, first, that the injected mRNAs are sufficiently stable that at neurulation the exogenous Xhox-1A mRNA still represents a 20-fold excess over the level of endogenous Xhox-1A, which was previously shown to peak at this time. They also showed that all three types of mRNA are translated well into protein and, in the case of Xhox-1A, the protein is transported into the nucleus. In contrast, the protein with a truncated homeobox remains in the cytoplasm, presumably because the signal for nuclear localization has been deleted. Finally, they also demonstrated that the injected mRNA diffuses from the site of injection enough to ensure that, following cleavage, most, if not all, of the daughter cells would receive some of the mRNA.

Drs. Harvey and Melton examined both the gross structure and the histology of embryos that developed from eggs injected with Xhox-1A mRNA. Fifty-one percent of the embryos injected with Xhox-1A mRNA showed a lateral kinking of their bodies after hatching. In contrast, only 13% of the embryos injected with beta-globin mRNA or mRNA from the homeobox-deleted Xhox-1A gene were defective, and the defects were variable. Histological examination after hatching of embryos injected with Xhox-1A mRNA showed that, while other embryonic tissues appeared normal, the highly ordered structure of the somites was completely lost. Somites are discrete blocks of tissue that form from the mesoderm (middle layer of the three embryonic tissue layers) adjacent to the notochord along the axis of all vertebrates. Somite formation begins in the anterior region of the embryo at neurulation and continues posteriorly until the late tailbud stage, when, in the frog, about 47 pairs of somites have formed.

Normally, somites give rise to organized muscle tissue. The observed results with exogenous Xhox-1A could have been due either to a disruption of pattern formation or to interference with differentiation of mesodermal precursor cells into muscle cells, although these two phenomena are related. To determine whether the cells in the abnormal somites had differentiated into muscle cells, Drs. Harvey and Melton used two muscle-specific probes. They found that cells in these abnormal structures in fact reacted with both an antisense-RNA for muscle-specific alpha-actin and with a muscle-specific monoclonal antibody. Based on their results, they concluded that the effect of the exogenous Xhox-1A is to disturb pattern formation and not cell differentiation.

Finally, they were able to show that endogenous Xhox-1A mRNA is expressed along the axis in the trunk region of the embryo, and predominantly in the somites. The Xhox-1A transcript is not found in the head or in cultured ectoderm. Thus, the site of expression of the Xhox-1A gene correlates with the site of the defect that they characterized in the injected embryos.

Their findings provide direct evidence that a homeobox-containing gene in Xenopus plays a role similar to that postulated for such genes in flies. The conclusion that homeobox-containing genes necessarily are involved in segmentation in all animals is premature, however. Evolutionary studies suggest that the common ancestors of vertebrates and arthropods (including flies) were not segmented. And unsegmented animals, such as echinoderms, also

have genes which contain homeoboxes. It is therefore possible that segmentation arose independently in vertebrates and arthropods and that homeoboxes were only later recruited in both evolutionary lines to assist in similar regulatory functions. Nevertheless, although the similarities between these genes in frogs and flies are circumstantial, the new findings from Dr. Melton's laboratory strongly suggest that vertebrate homeobox-containing genes are involved with pattern formation, in general, and with segmentation, in particular.

Judith H. Greenberg, Ph.D.

References

1. Harvey, R.P. and Melton, D.A. (1988). Microinjection of synthetic Xhox-1A homeobox mRNA disrupts somite formation in developing Xenopus embryos. Cell, 53, 687-697.

2. Harvey, R.P., Tabin, C.J., and Melton, D.A. (1988). Embryonic expression and nuclear localization of Xenopus homeobox (Xhox) gene products. EMBO J., 5, 1237-1244.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 27, 1989

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT - "Drosophila Homeobox Proteins are Transcription Factors" GM 37971 (Levine, Michael S.), Columbia University, and GM 37193 (O'Farrell, Patrick, H.), University of California, San Francisco

To Director, NIGMS

Establishment of the basic body structure of a Drosophila embryo depends on approximately 40 regulatory genes which are expressed in a highly time- and cell-specific manner. Based on studies of the timing and pattern of expression of these genes and the effects of mutations on embryonic development, the regulatory genes have been grouped into four classes which appear to interact with one another in a hierarchical manner. The gap genes are expressed first during embryogenesis, the pair-rule genes next, the segment polarity genes third, and the homeotic genes last. Over the past several years, considerable evidence has accumulated to suggest that, as a result of this complex hierarchy, virtually every cell is regulated by the activity of a unique set of genes.

Many of the regulatory genes contain a region that encodes a 60 amino acid sequence, called the homeobox, that is highly conserved in all multicellular animals. The deduced structure of the homeobox domain (a so-called helix-turn-helix) suggests that it is capable of binding to DNA. From this, it follows that proteins which contain homeoboxes are likely to be DNA-binding proteins. Consistent with this is the observation that all the homeobox-containing proteins that have been examined are located in the nucleus. Furthermore, several homeobox proteins which have been synthesized in E. coli from cloned DNA have been shown to bind in vitro to the 5' flanking regions of other genes in the regulatory network. Specifically, Dr. Patrick O'Farrell's group has shown that in the 5' flanking region of the engrailed (en) gene there is a series of DNA repeats with the consensus sequence TCAATTAAAT to which the en and fushi tarazu (ftz) homeobox domains bind in vitro (1). Dr. Michael Levine and his collaborators independently identified the same consensus sequence 5' to the en gene and showed that the homeobox proteins encoded by en, even-skipped (eve), zerknullt (zen), and paired (prd) bind to it in vitro (2). Although these results strongly suggest that homeobox proteins regulate transcription of other developmentally important genes, there has been no direct proof that they are transcription factors.

Dr. O'Farrell's laboratory (3) and a collaboration between the laboratories of Dr. Levine and Dr. James Manley (4) have now produced evidence that homeobox genes do indeed encode transcription factors which bind to specific DNA

sequences of genes further down in the developmental hierarchy. In addition, it appears that two or more homeobox proteins may compete or cooperate to regulate the expression of other developmental genes.

Both groups employed similar elegant strategies to answer the question of whether homeobox proteins regulate transcription in vivo. Each constructed two sets of plasmids. One set (producers) consists of expression vectors encoding the homeobox protein of interest. The second set of plasmids (responders) contains the target promoter (the presumed binding sites 5' of the en gene) fused to an easily assayable gene, such as chloramphenicol acetyl transferase (CAT) or beta-galactosidase. The producer and responder plasmids are introduced together into Drosophila tissue culture cells. Under appropriate conditions, the producer plasmid expresses the homeobox protein. If this protein is a transcriptional regulator of the en gene it should bind to the target en promoter on the responder plasmid and induce the synthesis of the marker protein (CAT or beta-galactosidase).

Dr. O'Farrell's group constructed responder plasmids containing a series of TCAATTAAAT consensus sequences as potential binding sites. They found that the ftz protein induced synthesis of the marker protein and that induction increased with increasing numbers of copies of the consensus site. Induction was observed when the consensus sequences were placed in either orientation and at varying distances from the marker gene. In contrast, when they introduced a producer plasmid which contained the en gene, they observed no induction of the marker gene in the responder plasmid. However, when producer plasmids for en and ftz were introduced into cells together, the expression of the marker protein was reduced compared to that observed with ftz alone. From these results they concluded that in vivo the ftz protein acts as a transcriptional regulator of en by binding to the same DNA sequences to which the protein binds in vitro. The en protein, which also binds to these same sites in vitro, probably competes with the ftz protein in vivo, repressing the activity of the en gene.

The group directed by Drs. Levine and Manley constructed responder plasmids containing approximately 2200 base pairs of DNA immediately 5' to the en gene, a region that includes six copies of the consensus binding sequence, and also plasmids containing various deletions in this 5'-flanking region. They then set out to analyze the effects of producer plasmids, which encoded various homeobox proteins (zen, prd, the zen-related gene z2, ftz, eve, and en), on expression of the marker protein. Producer plasmids which expressed zen, prd, and z2 induced expression of the marker protein in responder plasmids in which one of the six consensus sequences was deleted and the remaining five were juxtaposed to a position near the transcriptional start site. Surprisingly, however, when all six copies of the consensus sequence were intact, zen, prd, and z2 proteins did not induce significant synthesis of the marker protein.

This suggests that some portion of this region interferes with activation in this assay. The en and ftz producer plasmids did not induce expression of the marker gene in any of the responder constructs. Further experiments were done with responder plasmids with multiple repeats of a 96 base pair fragment, each containing three tandem copies of the consensus sequence. zen, z2, ftz, and prd all induced high levels of marker protein, whereas eve and en had no effect. In agreement with the observations of the O'Farrell laboratory, this group found that these repeats were activated in either orientation and if they were placed 3' to the marker gene.

The most interesting finding of the Levine-Manley group was that combinations of homeobox proteins act synergistically to induce expression of the marker protein. Specifically, plasmids that expressed the ftz, prd, and zen proteins alone each induced a relatively weak activation of transcription, but when producer plasmids for two or three of these proteins were introduced into the cells together, expression of the marker protein was significantly greater than predicted based on additivity. On the other hand, plasmids that expressed en or eve repressed activation induced by the z2 protein, and a plasmid that expressed eve inhibited activation induced by the ftz or prd proteins, or a combination of the two.

The results of both groups provide strong support for the idea that homeobox proteins are transcriptional regulatory factors in vivo. Futhermore, different homeobox proteins, which can bind to the same site on DNA, act together to modulate expression of a gene. This observation helps explain how developmental regulatory genes provide such exquisite control over early developmental events. It is likely that each cell in the developing embryo is exposed to a number of regulatory proteins at different times. DNA binding sites in each cell are probably occupied by mixtures of these proteins that vary in composition in a complex temporal and spatial pattern. As a result of competition for the same DNA binding site and of cooperation between two or more proteins in binding to the same site, virtually every cell might be subject to a unique regulatory influence. Both groups suggest that the transcriptional activity of the target gene depends on exactly what combination of homeobox proteins is bound.

Recently it was shown that some vertebrate transcription factors, including oct-1, oct-2, and GHF/Pit-1, contain conserved regions which resemble Drosophila homeoboxes (reviewed in 5). This discovery nicely complements the finding that Drosophila homeobox proteins are transcription factors. But it also injects a note of caution into the interpretation of the results described above. Specifically, if the homeobox motif is as common as it appears to be, it is likely that a large number of proteins would have the potential to bind to and transcriptionally activate many different genes, at least under experimental conditions. It is therefore possible that the examples of cooperation and competition between homeobox proteins observed in the cultured cells may not reflect the precise situation that exists in the developing embryo but rather just indicate some of the possible ways in which transcription factors might interact.

The widespread use of the homeobox as a DNA binding site also suggests that the specificity of binding of a homeobox protein must reside in regions of the protein outside the homeobox or must depend on interactions with other proteins.

As in the case of other fundamental discoveries, the identification of homeobox proteins as transcription factors in development raises as many new questions as it answers. It is clear that the next step in understanding developmental control will require experimental manipulations of the embryos themselves.

Judith H. Greenberg, Ph.D.

References

1. Desplan, C., Theis, J., and O'Farrell, P.H. (1988) The sequence specificity of homeodomain-DNA interaction. Cell 54:1081-1090.

2. Hoey, T. and Levine, M. (1988) Divergent homeo box proteins recognize similar DNA sequences in Drosophila. Nature 332:858-861.

3. Jaynes, J.B. and O'Farrell, P.H. (1988) Activation and repression of transcription by homeodomain-containing proteins that bind a common site. Nature 336:744-749.

4. Han, K., Levine, M.S., and Manley, J.L. (1989) Synergistic activation and repression of transcription by drosophila homeobox proteins. Cell 56:573-583.

5. Levine, M. and Hoey, T. (1988) Homeobox proteins as sequence-specific transcription factors. Cell 55:537-540.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 18, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Here, There and Everywhere: Regulation of the Yellow Gene in Drosophila" ROI GM 32036 (Corces, V.), Johns Hopkins University

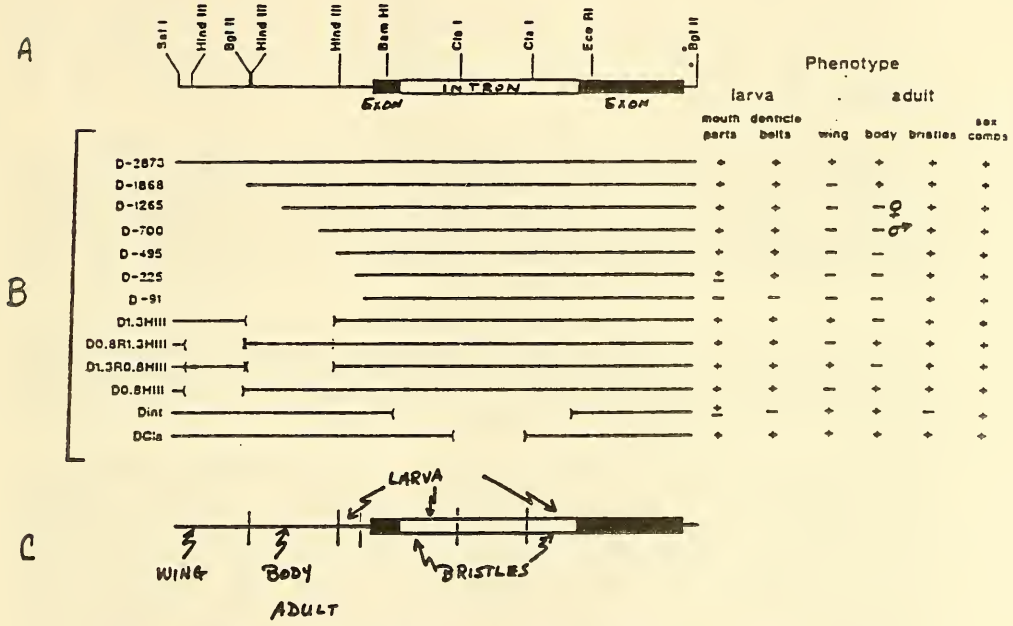
To Director, NIGMS
Through: Director, Genetics Program

Understanding how genes are regulated is a central problem for molecular geneticists, and NIGMS supports a large amount of research on this subject. For many years, grantee Victor Corces has been studying the yellow gene of Drosophila, a gene required for normal color of the body, wings, bristles and other structures (but not the eyes). As deduced from the DNA sequence, the gene product appears to be a secreted protein with a structural role in pigmentation, perhaps cross-linking melanin during cuticle tanning. More than twenty years ago, scientists identified two classes of yellow mutants: type 1's have defective pigmentation in all relevant structures whereas type 2's are deficient in some structures and not others. While type 1 mutants probably have mutations in the structural gene, type 2 mutants are thought to be regulatory mutants (expression of yellow has been shown to be regulated at the level of transcription, not translation). The few type 2 mutants analyzed molecularly have insertions in sequences upstream of (before) the transcribed region in potential regulatory regions.

Using now commonplace techniques, Dr. Corces and his postdoctoral associate Pamela Geyer (supported by NRSA fellowship F32 GM11156) isolated DNA containing the yellow gene and some of the upstream and downstream flanking DNA. A gene for eye color was attached to this piece of DNA as a selectable marker, and the engineered plasmid was inserted into the germ line of flies lacking yellow and having mutant eye color. Transformed flies with normal eye color presumably contained the plasmid with the yellow gene, and those containing only a single copy of the transforming DNA were selected for further study. DNA containing yellow plus 2873 base pairs of upstream DNA was able to restore wild-type coloration to transformed flies, showing that this plasmid contained all the necessary regulatory elements for proper expression of the gene. The scientists then made a series of transformation plasmids with varying amounts of upstream DNA deleted, a standard technique for identifying regulatory regions (see figure). Deleting the first 1000 bp from the 5' end led to loss of normal color in the wings. Removing another 600 bp led to loss of color in the body of females but not males; deleting the next 500 bp led to loss of body color in males as well.

These findings support the notion of control by several tissue-specific regulatory regions, with one for the wings being located far upstream and one (or more) for the body being closer to the coding region. An alternative hypothesis is that in different tissues, different amounts of gene product are necessary for producing color; correct wing color could require more protein and more transcription than body color, and in wings, the gene transcription would be responsive to more controlling elements.

To test this, the scientists removed the hypothetical body-color control region located after (3' to) the putative wing-color control region, which was left intact (plasmid D1.3HIII in figure). If the first hypothesis were correct, then this transforming DNA would affect only body color; if the second were true, then wing color would be affected. The transformed flies had normal wing color and mutant body color, thus supporting the hypothesis that in different tissues, gene expression is affected by different regulatory regions.



Plasmids used to identify enhancer elements controlling yellow. Figure A shows the sites within and around yellow where various restriction endonucleases cut DNA; these enzymes were used to make the plasmids of varying lengths shown in the Figure B. Spaces within parentheses indicate a region deleted within the plasmid; other deletions were made only from the 5' end. An arrow within the parentheses indicates that the correct sequence was inserted in reversed orientation. Figure C summarizes the locations of the enhancer elements.

Over the years, scientists have identified a number of different regulatory elements, including several that respond to exogenous factors or appear to be tissue specific. One of these is the enhancer element, unusual because it can function upstream, downstream, or even in the midst of the coding region, and can work in either orientation. To determine if the tissue-specific

regulatory regions for yellow were actually enhancer elements, Drs. Corces and Geyer constructed a plasmid in which the controlling region for either body color or wing color was placed in the opposite orientation; these constructs resulted in normal color in the appropriate tissues. These regulatory elements appear, therefore, to be enhancers that function independently to affect yellow expression in the two tissues.

The yellow gene contains a single large (2720 bp) intron, which is larger than the mRNA encoding the protein gene product. Since little is known about the function of introns, Drs. Corces and Geyer constructed plasmids with the intron deleted to determine if the intron affected transcription (see Highlight of March 30, 1988, "What's a nice intron like you doing in a gene like this?"). Flies transformed with this DNA had normal wing and body color but abnormal color in their bristles; however, if the transforming DNA had the ends of the intron intact but was missing DNA from the middle, color was normal in all body parts.

Interestingly, removal of the intron also affected yellow expression in larvae, and like bristle color in adults, the regulatory elements were located at the ends rather than in the middle of the intron. Whether the elements controlling adult bristle color and larval coloration are identical has not yet been determined. Another controlling region affecting larval coloration but not bristles or any other adult structure was found in a region very close (91 bp before) to the coding region.

Dr. Corces' results indicate that yellow has four, and probably more, enhancer elements affecting tissue-specific or developmental-specific gene expression. At least three are located in the typical control regions preceding the coding regions and one, or more, is located within the intron. Each seems to be used only in one tissue or during one stage of development.

These observations lead to the question now becoming paramount in studies of gene expression: how are the regulators regulated? Each enhancer element that affects yellow expression is itself controlled. A model of positive regulation would be for a specific protein to bind to the enhancer region, activating transcription; each enhancer element would have its own unique binding protein. What are these regulatory proteins? Are they present only at certain times or in certain tissues? What controls their production? Are these genes positively regulated? Could differential gene regulation be negative regulation, that is, is a gene turned on unless specifically turned off? As we learn more about gene regulation, we can become frustrated and awed by its complexity. But such precise and exquisite control must be complex because errors would be disastrous in such a fundamental process.

Reference:

Geyer, P. K. and V. G. Corces. (1987). Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in Drosophila melanogaster. Genes & Development 1:996-1004.

Joye F. Jones, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date October 24, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "MyoD1: A Master Regulator of Myogenesis"; RO1 GM26176, (Weintraub, H.) Fred Hutchinson Cancer Research Center

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 10-24-89

Our current understanding of how transcription is regulated is based upon a simple model in which transactivating protein factors bind to specific regions of DNA around a gene to "turn on" expression of that gene. In most cases, a number of different factors are required to activate transcription. Some of these are ubiquitous and are required for the expression of many genes; others are highly specialized factors present only in certain tissues. The exact stoichiometry and relative abundance of these factors in any given tissue leads to the tissue-specific expression of certain genes. This type of model has been well documented for many different cell types. Examples of well-studied systems are the expression of the globin family of proteins in erythrocytes or of myosin heavy and light chains in muscle cells.

Superimposed on this local regulatory mechanism must be a higher order level of organization required for the coordinated expression of a whole class of structural genes associated with a differentiated phenotype. The simplest model for this type of control would be one in which the product of a single master regulatory gene could turn on an entire program of differentiation. Based on what is already known about how transcriptional regulation occurs at the local level, the predicted mechanism would be one in which a single protein factor binds to a region in a structural gene to "tag" it for expression in a specific cell lineage. The coordinated expression of the tagged genes would in turn lead to the new phenotype. Now there is evidence that such a factor actually exists. Dr. Harold Weintraub and his colleagues at the Fred Hutchinson Cancer Research Center have shown, in an elegant series of papers published over the past three years, that a single gene product, called MyoD1, acts as a master switch which converts nonmuscle cells to myoblasts.

The first indication that a master regulatory protein might exist came from a series of experiments designed to study the relationhip between DNA methylation and gene expression. It had been known for some time that demethylation of DNA can activate gene expression. This effect can be studied by using 5-azacytidine to inhibit the methylation of cytosine residues. Three years ago, Dr. Weintraub showed that brief exposure of mouse fibroblasts, called 10T1/2 cells, to 5-azacytidine would convert these cells to stable myoblasts. Two sets of gene transfer experiments were then devised to establish that the conversion was due to the activation of a "master gene" controlling the downstream events associated with muscle differentiation. In the first set of experiments, a demethylated cardiac muscle actin gene was transfected into "azamyoblasts" (10T1/2 cells converted to myoblasts by 5-azacytidine treatment) and into untreated 10T1/2 cells. The cardiac actin gene was expressed at high levels only in the azamyoblasts. Analysis of the RNAs expressed in these cells showed that the

endogenous skeletal and cardiac actins were also expressed. Since transfer of the cardiac muscle gene by itself was not enough to induce its own expression, this result was taken to mean that demethylation had resulted in activation of a regulatory gene responsible for induction of myoblast-specific structural genes. In the second DNA-mediated gene transfer experiment, DNA isolated from azamyoblast cells and untreated 10T1/2 cells was transfected into untreated 10T1/2 cells. Expression of muscle-specific genes was found only among those transformants treated with the azamyoblast DNA, again suggesting that a regulatory gene had been transferred which induced the expression of a repertoire of structural genes. The frequency at which colonies could be converted from fibroblasts to myoblasts was consistent with the transfer of one or only very few closely linked demethylated genes (1).

Characterization of the precise locus involved in converting 10T1/2 cells to determined myoblasts came from the next set of experiments in which a cDNA clone carrying the sequence responsible for the conversion activity was identified (2). This cDNA clone was identified by screening a myocyte cDNA library with mRNAs expressed in myoblasts but not in 10T1/2 cells. The cDNA clone that was selected, called MyoD1 (Myoblast Determination Gene Number 1), was then cloned into an expression vector and transfected into 10T1/2 cells. Expression of the gene resulted in synthesis of the myosin heavy chain and a phenotype morphologically indistinguishable from that of myoblasts. Thus, the effect of transferring one gene to a nonmuscle cell was enough to turn on expression of other genes associated with the muscle phenotype.

A survey of various tissues demonstrated that MyoD1 mRNA is normally present only in neonatal and adult skeletal muscle tissue, not in smooth muscle or in any other tissue tested, including liver, kidney, lung, spleen, brain, or ovary. MyoD1 is found in the nucleus. It is a phosphoprotein of 318 amino acids which includes a short segment that is homologous to a sequence present in the *myc* family of proteins (2,4). This short amino acid sequence in *myc* is within a region that is important for oncogenic transformation. The same region of *myc* homology has been found in several Drosophila genes known to be involved in development, including *achaete-scute* (neurogenesis), *daughterless* (sex determination), and *twist* (mesoderm formation). Further analysis of the functional domains of MyoD1 indicate the presence of a highly basic region that is required for both nuclear localization and induction of myogenesis. Deletion of the short region of *myc* homology eliminates the ability of MyoD1 to initiate myogenesis, but does not alter nuclear localization. Expression of only 68 amino acids containing the basic and the *myc* domains is sufficient to activate myogenesis (4).

Although it was clear that MyoD1 could convert a variety of stable murine fibroblast lines to muscle, the effect of MyoD1 on fibroblast lines from different species, on primary, nontransformed cells and on differentiated tissue culture lines was not known. Dr. Weintraub and coworkers tested the effect of transfecting a MyoD1 cDNA expression vector on primary chick, human and rat fibroblasts. Not only did these cells show positive staining for the muscle-specific markers myosin and desmin, but they became elongated, began to fuse and formed large multinucleated myotubes (3). Muscle markers could also be activated by MyoD1 in differentiated melanoma, neuroblastoma, liver, kidney, and adipocyte

tissue culture lines. These findings demonstrate that MyoD1 can induce muscle differentiation across species, in both transformed and primary cell lines, and in cells derived from each of the three (mesodermal, ectodermal and endodermal) germ layers. It is also clear from these experiments that no factors other than MyoD1 are required to activate the downstream program for terminal muscle differentiation (3).

An analysis of the mechanism of MyoD1 activation has shown that transfection of 10T1/2 cells with MyoD1 cDNA induces expression of endogenous MyoD1 mRNA, indicating that MyoD1 is subject to positive autoregulation (5). Surprisingly, transfection with MyoD1 also activates expression of a second muscle-specific protein, myogenin, which can regulate MyoD1 and autoregulate itself. During normal development, myogenin appears in a temporal pattern of expression which is different from that of MyoD1. While MyoD1 is expressed in myoblasts and in skeletal muscle, myogenin is absent from myoblasts, peaks during myofiber formation, and then declines in well differentiated myotubes. Although a distinctly different protein, myogenin shares with MyoD1 the amino acid region that is structurally related to the *myc* genes. Thus, MyoD1 and myogenin appear to be involved in a positive autoregulatory loop which might function to maintain or amplify their expression in order to stabilize the myogenic commitment, each acting to initiate a different type of muscle lineage during development (5).

Muscle cell differentiation entails both withdrawal from the cell cycle and activation of a repertoire of muscle-specific differentiation markers, including actin, myosin heavy and light chains and acetylcholine receptor. As in many other cell lineages that terminally differentiate, cellular proliferation and synthesis of differentiation markers are mutually exclusive processes in myogenic cells. The expression of transforming oncogenes in muscle cells inhibits cellular differentiation. In muscle cells, expression of oncogenic tyrosine kinases (v-*src*, v-*fps*), growth factor receptors (v-*erbB*), nuclear oncogenes (v-*myc*, c-*myc*, v-*erbA*, and *E1A*), and the activated form of putative signal transducing G proteins (H-*ras*, N-*ras*) can inhibit terminal differentiation to varying extents. It is unclear whether oncogene products abort differentiation by inhibiting withdrawal from the cell cycle, for example, by providing a constant mitogenic signal, or if they more directly interfere with expression of differentiation-specific genes. In one of his recent publications, Dr. Weintraub demonstrates that activated *ras* inhibits not only muscle-specific markers of terminal differentiation but also expression of the myoblast lineage markers MyoD1 and myogenin in transfomed myoblasts (6). Expression of MyoD1 in these *ras*-transformed cells causes the re-expression of both terminal differentiation markers and the lineage markers, including endogenous MyoD1 and myogenin, even though activated *ras* is still present. Hence, MyoD1 overrides the *ras*-induced inhibition of differentiation, suggesting that *ras* might act by silencing regulatory loci, such as MyoD1.

In his latest paper on MyoD1, Dr. Weintraub and colleagues have shown that MyoD1 is a sequence-specific DNA binding protein capable of specific interaction with the enhancer region of a muscle-specific structural gene, creatine kinase (7). The enhancer region is known to be required for full activity of the creatine kinase gene. The basic and *myc* homology regions that were previously shown to be required to convert 10T1/2 fibroblasts into myoblasts are the same regions

of the protein required for specific protein-DNA interaction. Thus, it is likely that the biological activity of MyoD1 is mediated through its capacity for specific DNA binding. In turn, the mechanism for overall induction of muscle differentiation appears simply to be the expression of a single determination factor that can specify muscle cell lineage by binding to the regulatory regions of tissue-specific structural genes. The discovery of a master switch protein such as MyoD1 is a fundamental step in understanding how the process of differentiation is initiated and maintained in a developing system.

Catherine Lewis

Catherine D. Lewis, Ph.D.

References

1. Lassar, A.B. Paterson, B.M., and Weintraub, H. (1986) Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. Cell 47: 649-656.

2. Davis, R.L., Weintraub, H., and Lassar, A.B. (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987-1000.

3. Weintraub, H., Tapscott, S.J., Davis, R. L., Thayer, M.J., Adam, M.A., Lassar, A.B. and Miller, D. (1989) Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. PNAS 86: 5434-5438.

4. Tapscott, S.J., Davis, R.L., Thayer, M.J., Cheng, P.-F., Weintraub, H. and Lassar, H. (1988) MyoD1: A nuclear phosphoprotein requiring a *myc* homology region to convert fibroblasts to myoblasts. Science 242: 405-411.

5. Thayer, M.J., Tapscott, S.J., Davis, R.L., Wright, W.E., Lassar, A.B. and Weintraub, H. (1989) Positive autoregulation of the myogenic determination gene MyoD1. Cell 58: 241-248.

6. Lassar, A., Thayer, M., Overell, R. and Weintraub, H. (1989) Transformation by activated ras or fos prevents myogenesis by inhibiting expression of MyoD1. Cell 58: 659-667.

7. Lassar, A.B., Buskin, J.N., Lockshon, D., Davis, R.L., Apone, S., Hauschka, S.D., and Weintraub, H. (1989) MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. Cell 58: 823-831.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 9, 1989

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Cell Cycle Control. I. Identification of a Cytoplasmic Regulator of Mitosis" GM 34607 (David Beach), Cold Spring Harbor Laboratory; GM 33523 (John Newport), University of California, San Diego; GM 26743 (James Maller), University of Colorado*

To Director, NIGMS

As a model system, yeast offers a tremendous advantage over other eukaryotic organisms because of the ease with which its genetics can be manipulated. However, yeast is usually not the best system in which to do biochemical studies. For this reason, recent collaborations between laboratories interested in cell cycle control in yeast and in induction of mitosis in Xenopus eggs, has paid off in a big way.

Over the last 15 years a variety of mutations that perturb the normal progression through the cell cycle have been isolated in both the budding yeast, Saccharomyces cerevisiae, and in the fission yeast, Schizosaccharomyces pombe, suggesting that a number of genes are required to regulate the cell cycle. In S. pombe, one of the key cell cycle control genes is cdc2, which encodes a 34 kilodalton (kd) protein. The cdc2 gene product acts early in the cell cycle (at the G1-S transition), where it is required for the cell to become committed to undergo mitosis and to initiate DNA synthesis, and also acts later in the cell cycle (at the G2-M transition) to trigger mitosis. (In S. cerevisiae, a gene called CDC28, encodes a homologous protein which plays a similar role.) Paul Nurse in Oxford, England, and David Beach, at that time working in the Nurse laboratory, found that the cdc2 protein is a protein kinase, an enzyme which adds phosphate groups to specific sites on proteins (1). Several other genes whose products interact with the cdc2 protein to regulate mitosis were later identified, including suc1, which encodes a 13 kd protein. However, because cell-free assays for mitotic induction have not been developed in yeast, it has not been possible to study biochemically the roles of the various cell cycle control proteins that are assumed to exist.

Studies on mitosis in Xenopus oocytes have an equally long history but have followed a very different path. An unpurified factor, known as maturation promoting factor (MPF), has been known since the early 1970s to induce mitosis in Xenopus eggs and was subsequently found to induce mitosis in somatic cells of all eukaryotic organisms tested. In recent years, the laboratories of NIGMS grantees James Maller and John Newport, among others, have developed

*Dr. Maller's grant is in the CMBD Program. Cell cycle control is an example of a research area in which genetics and cell biology have made a fruitful union.

cell-free systems in which many of the events in the cell cycle can be mimicked by adding appropriate factors to reconstituted nuclei in the presence of interphase cytoplasm. For example, when MPF is added to interphase cytoplasm, the nuclear envelope surrounding the reconstituted nuclei breaks down, an early step in the MPF-dependent entry into the mitotic state. Using this cell-free system as an assay, members of Dr. Maller's laboratory identified and purified two proteins, with molecular weights of 32 and 45 kd, which are the active components of MPF. These proteins have protein kinase activity (2).

The fact that one of the active factors of Xenopus MPF and the yeast cdc2 protein have similar molecular weights and are protein kinases set the stage for two separate collaborative efforts, one between the laboratories of Drs. Newport and Beach and the other between the laboratories of Drs. Maller and Nurse. Using different approaches, the two groups have recently demonstrated that purified MPF from Xenopus and the cdc2 protein in S. pombe are very closely related and probably have similar functions (3,4).

The Maller and Nurse group (3) found that antibodies against the S. pombe cdc2 protein reacted with a 34 kd protein in Xenopus egg extracts. Furthermore, the cdc2 antibodies immunoprecipitated both a 32-34 kd protein and a 45 kd protein from their purified MPF, suggesting that the 34 kd protein can exist in a complex with the 45 kd protein. The precise role of the 45 kd protein is not known, but these investigators suggest that it might be needed to activate or regulate the activity of the 34 kd protein.

Dr. Beach's group, in earlier studies, had found that overproduction of the suc1 protein in S. pombe causes a delay in cell division (5). The Beach and Newport group (4) therefore used the cell-free system described above to test the effect of the yeast suc1 protein on nuclear envelope breakdown as an indication of whether it affects MPF-induced initiation of mitosis. They discovered that the suc1 protein inhibited MPF-induced mitosis in the cell-free system. This finding has biological significance because the suc1 protein acts at low concentrations, is not a general inhibitor of nuclear events, and its action can be reversed by adding more MPF. The yeast suc1 protein, which is known to interact with the yeast cdc2 protein, was also shown to associate physically with the fraction of Xenopus MPF that has a molecular weight of 33-34 kd. Furthermore, the yeast cdc2 protein and the MPF activity have identical chromatographic properties on a suc1 protein affinity column. The only other protein from crude Xenopus MPF that bound to the suc1 protein affinity column was a 42 kd protein, which is probably identical to the 45 kd protein purified by the Maller and Nurse group. Taken together, the work of the two groups provides strong evidence that the yeast cdc2 protein is the same as Xenopus MPF or is a component of MPF necessary for its mitosis-inducing activity.

A body of circumstantial evidence suggests that protein phosphorylation is an important aspect of mitotic control. It is tempting to speculate that the

yeast cdc2 protein, which has protein kinase activity, and its MPF homolog in Xenopus represent the postulated pleiotropic stimulator of a protein kinase cascade that would initiate the morphological alterations that occur early in mitosis.

It is of interest that, in S. pombe, the amounts of cdc2 protein and its kinase activity do not fluctuate during the cell cycle, and mutant yeast cells in which cdc2 protein is overproduced divide normally. This suggests that another protein is required for the proper functioning of cdc2 protein in vivo. The obvious candidate is the 42-45 kd protein identified by both groups. Thus, active MPF might consist of a complex between the cdc2 protein and the 42 kd protein.

While much remains to be learned about the role of cdc2 protein-MPF, the recent work described here provides a link between yeast and amphibian mitotic regulatory systems that will enable biochemical and genetic approaches to be combined. Proteins associated with MPF which are potential substrates in Xenopus can be compared with gene products thought to interact with the cdc2 protein in yeast. These studies reinforce the view that the basic mechanism of mitotic initiation is similar in all eukaryotes and give new momentum to the very exciting and fast-moving research underway in cell cycle control.

Judith H. Greenberg, Ph.D.

References

1. Beach, D.H., Durkacz, B., and Nurse, P.M. (1982). Functionally homologous cell cycle control genes in budding and fission yeast. Nature 300:706-709.

2. Lohka, M.J., Hayes, M.K., and Maller, J.L. (1988) Purification of maturation-promoting factor, an intracellular regulator of early mitotic events. Proc. Natl. Acad. Sci. USA 85:3009-3013.

3. Gautier, J., Norbury, C., Lohka, M., Nurse, P., and Maller, J. (1988) Purified maturation-promoting factor contains the product of a Xenopus homolog of the fission yeast cell cycle control gene cdc2+. Cell 54:433-439.

4. Dunphy, W.G., Brizuela, L., Beach, D., and Newport, J. (1988). The Xenopus cdc2 protein is a component of MPF, a cytoplasmic regulator of mitosis. Cell 54:423-431.

5. Hindley, J., Phear, G., Stein, M., and Beach, D. (1987) Suc1+ encodes a predicted 13-kilodalton protein that is essential for cell viability and is directly involved in the division cycle of Schizosaccharomyces pombe. Molec. Cell. Biol. 7:504-511.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 9, 1989

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Cell Cycle Control. II. The Human Homolog of a Yeast Gene Encodes a Major Cell Cycle-Regulated Tyrosine Kinase Substrate" GM 34607 and GM 39620 (David Beach), Cold Spring Harbor Laboratory

To Director, NIGMS

The factors that regulate the cell division cycle in mammalian cells are poorly understood, in part because of the difficulty in identifying critical cell cycle regulatory proteins in organisms that are not amenable to classical genetic techniques. As a result, for genetic studies of the eukaryotic cell cycle, scientists have turned to more tractable models -- the budding yeast, Saccharomyces cerevisiae, and the fission yeast, Schizosaccharomyces pombe.

Two major controls points operate during the cell cycle in S. pombe. The first occurs in late G1, just before DNA synthesis begins. Once past this point, cells are committed to undergo mitosis rather than to follow the pathway leading to conjugation. The second control occurs late in G2 and determines when mitosis is initiated. Of the many genes specifically required for progression through the cell cycle, the cdc2 gene is the only one known to be essential at both control points. The cdc2 gene (and its homolog, CDC28, which plays a similar role in S. cerevisiae) encodes a 34 kilodalton (kd) protein, designated p34, which has protein kinase activity. (Protein kinases catalyze the addition of phosphate groups to specific sites on proteins.) A 13 kd protein, p13, encoded by the suc1 gene, forms a stable complex with the cdc2 protein kinase and is essential for cell cycle progression.

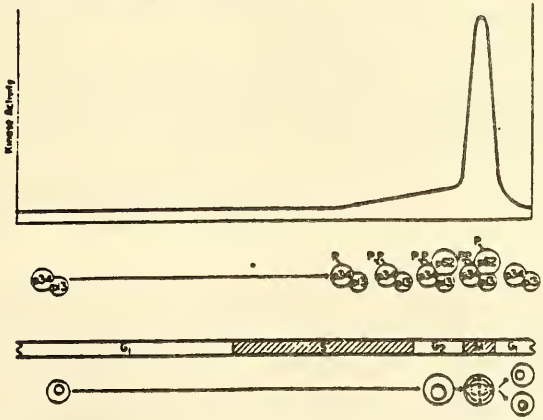
Given the central role of cdc2 in the cell cycle of S. pombe and of CDC28 in that of S. cerevisiae, it was of interest to establish whether a similar gene function exists in mammalian cells. David Beach and members of his laboratory found that antibodies prepared against the yeast cdc2 protein immunoprecipitate a 34 kd protein from human HeLa cells (1). Like the yeast cdc2 protein, the human p34 is a protein kinase and it exists in a complex with a 13 kd protein homologous to the yeast suc1 protein. Furthermore, biochemical analysis showed that the yeast and human p34 proteins are structurally related. The similarity between the yeast and human genes, however, does not necessarily imply that the biological roles of the two proteins are identical. Dr. Beach's group therefore set out to investigate the properties of the human p34 protein kinase in HeLa cells and its activity at different stages of the cell cycle (2).

By immunoprecipitation of extracts of non-synchronized HeLa cells, this time with antibodies against human p34, they found first that human p34 is composed of three species that are resolved by electrophoresis. The one that migrates most rapidly has an electrophoretic mobility identical to that of yeast p34;

the other two species were found to be phosphorylated forms of the protein. Several other proteins were also immunoprecipitated, two of which exist in a complex with the most highly phosphorylated form of p34. One of these has a molecular weight of 13 kd and corresponds to the previously identified sucl protein homolog. The second has a molecular weight of 62kd and does not correspond to any known yeast gene product. Further experiments showed that when p34 is associated with p13, the complex can phosphorylate casein, a standard substrate for protein kinases, but when p34 is associated with p62, the complex not only phosphorylates casein but also p62 itself; unphosphorylated p34 does not have significant kinase activity.

Dr. Beach found that the total abundance of p34 varies little during the cell cycle, but that its protein kinase activity, with respect to both casein and p62 as substrates, increases greatly at the late G2-M transition and reaches a peak during mitotic metaphase. From these results, he concluded that the state of phosphorylation of p34, its association with p62, and the protein kinase activity of the p34-p62 complex are all subject to cell cycle regulation. Thus, in G1, p34 is unphosphorylated, unassociated with p62, and inactive as a kinase. In contrast, at the transition from G1 to M, p34 is phosphorylated, associated with p62, and active as a kinase.

Dr. Beach has postulated that activation of the p34 complex is a multistep process (figure). Throughout the cell cycle, p34 is associated with p13. During G2, p34 becomes phosphorylated at two sites, and at the G2/M transition it associates with p62. These investigators believe that there is an additional, unidentified step that further activates the complex. At the end of mitosis, the steps are reversed.



Relationship between the cell division cycle and the proposed cycle of p34 kinase. p34 is associated with p13. During the cell cycle, p34 becomes phosphorylated at two sites, associates with p62, and is finally activated by an unidentified event during the transition to mitosis. These steps are reversed before cytokinesis. Direct contact between p34, p13, and p62 are shown only for simplicity (from ref. 2).

Phosphorylation of proteins can occur at serine and threonine residues and, less commonly, at tyrosine residues. Dr. Beach's group found that p34 contains predominantly phosphothreonine, but also a significant amount of phosphotyrosine (3). In fact, p34 is the most abundant phosphotyrosine-containing protein in HeLa cells. Moreover, the level of tyrosine phosphorylation, like that of total phosphorylation measured in the earlier experiment, is regulated during the cell cycle.

Among the known tyrosine kinases, only one, the product of the pp60c-src gene, has been shown to be subject to cell-cycle regulation. (c-src is the normal cellular counterpart to the viral oncogene, v-src.) Dr. Beach's group found that pp60c-src phosphorylates p34 in vitro at a single site. In vivo, p34 is phosphorylated at the identical site and also at two other tyrosine residues. These observations indicate that either pp60c-src or a similar tyrosine kinase regulates p34 activity in vivo, but since only one of the three phosphorylated sites observed in vivo is also phosphorylated in vitro, p34 is probably the substrate of more than one tyrosine kinase.

Taken together, these experiments show that human p34 is both a protein kinase and a substrate for protein kinases. Thus, it may be part of a pathway analogous to known growth factor signal transduction pathways. These results also raise interesting questions about the role of oncogenes in the regulation of cell division. In yeast and, based on Dr. Beach's work, also in higher eukaryotes, cdc2 regulates passage through the cell cycle, but there is no evidence that it is involved in growth control. Nevertheless, it should not be surprising that p34, the cdc2 gene product, might be a substrate for pp60c-src, since one element of oncogenesis must be the activation of cell-cycle pathways. It might be relevant that the viral counterpart of pp60c-src, when tested in other systems, has a similar substrate specificity but has greater tyrosine kinase activity.

Dr. Beach's research is a prime example of how studies on a simple model system can lead to profound new understanding of normal human physiology, and possibly even of carcinogenesis.

Judith H. Greenberg, Ph.D.

References

1. Draetta, G., Brizuela, L., Potashkin, J., and Beach, D. (1987) Identification of p34 and p13, human homologs of the cell cycle regulators of fission yeast encoded by cdc2+ and suc1+. Cell 50:319-325.

2. Draetta, G. and Beach, D. (1988) Activation of cdc2 protein kinase during mitosis in human cells: Cell cycle-dependent phosphorylation and subunit rearrangement. Cell 54:17-26.

3. Draetta, G., Piwnica-Worms, H., Morrison, D., Druker, B., Roberts, T., and Beach, D. (1988) Human cdc2 protein kinase is a major cell-cycle regulated tyrosine kinase substrate. Nature 336:738-744.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date August 18, 1988

From Program Administrator, Genetics Program, NIGMS

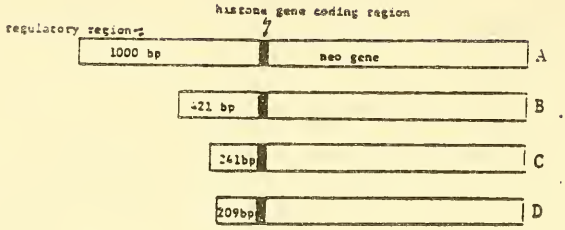
Subject RESEARCH HIGHLIGHT: "Histone Gene Control: A Model for Cell-cycle Regulated Gene Expression" RO1 GM 31138 (Lee, A.), University of Southern California, Los Angeles

To Director, NIGMS
Through: Director, Genetics Program, NIGMS

Scientists in many laboratories are working to develop a comprehensive explanation of the extremely complex process of gene regulation. Certain regions of DNA immediately preceding (upstream of) the region encoding the protein product are critical for gene transcription. In eukaryotes, these DNA sequences include the promoter region to which RNA polymerase binds as well as the CAAT and TATA "boxes", conserved DNA sequences that are often necessary for efficient transcription. Much less is known about spatial or temporal regulation of genes, genes transcribed only in certain tissues, at certain times, or in response to external factors. Studies on the effects of heat shock or hormones are yielding important information on how genes are affected by these factors. Dr. Amy Lee is studying another aspect of gene regulation, temporal regulation, and in particular, the effect of the cell cycle on gene activation.

Although a number of genes are activated during specific phases of the cell cycle, the most studied are the genes involved in the production of histones, the basic proteins that bind to DNA in chromatin. There are different kinds of histones, and in many organisms, the multiple histone genes are arranged in tandem clusters. The entire cell complement of histone is made during the relatively short period of DNA replication, the S phase. Because the histone genes are most actively transcribed during S phase, with only very low levels of transcription at other times, they represent an excellent system in which to study cell-cycle control of gene expression.

Dr. Lee and her colleagues, including Elpidio Resendez, a predoctoral fellow supported by an individual fellowship (F31 GM 09805) from the MARC Program, have been trying to identify the nucleotide sequences involved in cell-phase-specific regulation of histone gene expression. Because many regulatory sequences are found upstream of the coding region, they concentrated on the DNA in this area. Since measuring histone production directly is a cumbersome process, Dr. Lee's group used an engineered gene that was easier to study. The potential regulatory region, consisting of 1000 nucleotides of DNA from the region upstream of a hamster histone gene, plus 60 nucleotides of the histone gene coding region itself, was attached to a selectable marker gene (called neo) encoding resistance to the drug neomycin (A in the figure). Only cells expressing the neo gene under the control of the histone gene regulator would grow in the presence of an analog of the drug. When inserted into hamster kidney cells growing in tissue culture, this engineered neo gene was shown to be under cell-cycle control, being expressed most actively during S phase and at significantly lower levels during G_1 and G_2 phases.



Four of the engineered genes used to identify the critical control region.

To locate the cell-cycle control region, the group in Dr. Lee's laboratory constructed progressively shorter variants of their artificial gene by removing nucleotides from the 5' end of the potential regulatory region (see figure). They left intact the nucleotides closest to the coding sequences that would contain the promoter region and CAAT and TATA boxes. With only 241 nucleotides still present (C in the figure), the neo gene was expressed fully and remained under cell cycle control. Removing another 32 nucleotides (D in the figure) reduced the level of transcription to that found in G_1 cells, and the gene was no longer controlled by the cell cycle. This critical 32 nucleotide region contained a six-nucleotide sequence motif similar to one found in two yeast genes known to be under cell cycle control, and similar also to the regulatory region of an analogous murine histone gene. The regulatory region identified by Dr. Lee may function as an enhancer element, but it is much smaller than other well-described enhancers, which are often hundreds of base pairs long and usually contain repeated sequences.

One way of activating this gene in a cell-cycle-specific manner would be for a protein, present only during certain phases of the cell cycle, to bind to this regulatory region. Dr. Lee and colleagues found two proteins that would bind specifically to the histone regulatory element. One protein was present during all phases and presumably accounts for the low level of histone production throughout the cell cycle. The second was present in highest concentrations during the end of G_1 and the beginning of S, exactly the time a regulatory protein should be active. The job now is to identify the gene which produces this protein--it is not histone--and determine what regulates its expression. Dr. Lee has preliminary evidence, not yet confirmed, that the regulatory protein is the product of a proto-oncogene; if true, this would add to the growing evidence that proto-oncogenes commonly regulate cell growth.

We are beginning to understand fairly well how steroidal hormones affect gene expression, but we have a long way to go before we understand how other factors work. This research is particularly interesting because Dr. Lee has been able to find a regulatory region used during one phase of the cell cycle. With the cell-cycle specific control region identified, scientists have a starting point for understanding how genes respond to the cell cycle.

Joye P. Jones, Ph.D.

Reference:

Artishevsky, A., S. Wooden, A. Sharma, E. Resendez, Jr., and A. S. Lee. 1987. Cell-cycle regulatory sequences in a hamster histone promoter and their interactions with cellular factors. Nature 328:823-827.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 22, 1989

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "It's Hard to Splice When You're Hot" R01 GM25874 (Lindquist, Susan), University of Chicago

To Director, NIGMS
Through: Director, Genetics Program, NIGMS 9-22-89

Cells respond to high temperatures by synthesizing heat shock proteins and suppressing the synthesis of most other normal cellular proteins. Both transcription and translation are affected, and during heat shock, only heat shock genes are transcribed into messenger RNA. Heat shock mRNA's are translated very rapidly, whereas pre-existing mRNA's are neither translated nor degraded. Instead, they remain unaffected during the heat treatment and are then translated during recovery from heat shock. When the temperature returns to normal, heat shock mRNA's and proteins are degraded very rapidly. The heat shock response is universal, and both the genes and the heat shock proteins (hsp's) they encode have highly conserved sequences. Finally, the heat shock response appears to be a generalized stress response since it can be induced by a variety of factors other than heat, including ethanol, damaged proteins, oxidizing agents, and DNA damaging agents.

NIGMS grantee Susan Lindquist has been studying the heat shock response in the fruit fly, Drosophila melanogaster. Drosophila cells are usually cultured at 25°C, approximately normal room temperature. The heat shock response can be induced by shifting the cells to 29-39°C, and is maximal at 37°C. For a number of years, Dr. Lindquist and her graduate student Joseph Yost (supported by NIGMS predoctoral genetics training grant, T32 GM07197) have been studying another effect of heat shock on protein synthesis, heat shock inhibition of RNA splicing. Most eukaryotic genes contain regions encoding the protein gene product which are interrupted by regions of DNA that seem to be nonsense. The coding sequences are called exons, and the intervening sequences, introns. The function of introns is unknown. (See Research Highlight of March 30, 1988, "What's a nice intron like you doing in a gene like this?".) Genes containing introns are transcribed as a single piece of precursor mRNA; the introns are subsequently cut out and the exons are spliced together to form the mature mRNA that is translated into protein.

Heat shock genes in Drosophila are unusual in that most of them do not contain introns. However, one heat shock gene, called hsp83, does contain an intron, and the effect of heat on hsp83 production is quite different from its effect on other hsp's. Although abundant at intermediate temperatures (33-35°C), hsp83 is produced at only very low levels in cells heat shocked at high temperatures (37-38°C). If, however, cells are incubated first at the intermediate temperature (where they produce large amounts of hsp83) and then at the high one, the cells continue to produce large amounts of hsp83. This phenomenon is not unique to the hsp83 gene, because a specially constructed artificial gene behaved the same way. This engineered hsp-adh gene had the promoter region from the hsp70 gene connected to the structural gene (which contains introns) for alcohol dehydrogenase (adh). (The scientists had to use a heat shock gene promoter so that the engineered gene would be transcribed at high temperatures since the adh gene with its normal promoter does not function during heat shock.) The hsp83 and hsp-adh genes were transcribed at a higher rate at 38°C, a temperature that completely inhibits protein synthesis, than at 33°C, when protein synthesis is greatest. Protein was not produced at the high temperature because precursor mRNA was not processed into mature mRNA. In fact, the process of cutting and splicing did not even

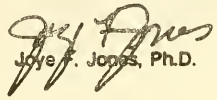
begin; the precursor mRNA was not cut at all. Because heat shock genes other than hsp83 do not have introns, their mRNA does not have to be processed in order to be translated.

Although hsp83 and adh were not made at the higher (38°C) temperature, the ribosomes did begin translating unprocessed mRNA. Drs. Yost and Lindquist showed this by engineering another hsp-adh gene that had a "stop" signal within the first intron. If the RNA were translated correctly, that is, if the ribosomes "skipped over" the intron, the result would be a 50kD polypeptide. If the ribosome read the RNA incorrectly and attempted to translate the intron, it would soon reach the stop sign; the result would be a 40kD polypeptide. At a low temperature that permits mRNA processing, ribosomes made the 50kD polypeptide. At the higher inhibitory temperature, they made only the 40kD polypeptide. Thus, at the higher temperature, the ribosome was reading into the intron, encountering the stop codon, and terminating the protein prematurely. It is easy to see that if all genes were transcribed during heat shock, the cell would soon be full of damaged and non-functional proteins.

Cells were able to make large amounts of hsp83 and adh at a high temperature (38°C) if they were incubated first at an intermediate temperature (35°C). However, if the cells were prevented from making protein (by adding a inhibitor of protein synthesis) at the intermediate temperature, they were subsequently unable to make protein when shifted to the higher temperature. The cells could transcribe hsp83 and hsp-adh genes normally at both temperatures, and they continued to produce the other hsp's at normal levels; only the ability to process the RNA of intron-containing genes was affected. The ability to make protein during the 35°C incubation, therefore, is necessary for heat shock protection of splicing, and the predominant proteins made at this temperature are heat shock proteins. Heat shock proteins must, therefore, be necessary for proper cutting and splicing of precursor mRNA during heat shock.

The heat shock response is truly fascinating. The heat shock genes appear to have been under strong evolutionary constraints since they are among the few eukaryotic genes lacking introns. Interestingly, the only other genes transcribed to any significant degree during heat shock are histone genes; they also lack introns. One must wonder why hsp83 has an intron. Perhaps it has some other primary function and only fortuitously has a structure that, in the right circumstances, allows it to be transcribed and translated at high temperature.

The sequence of events that make up the heat shock response clearly help the cell survive otherwise damaging temperatures. RNA processing appears to be very sensitive to heat. By turning off transcription of non-heat shock genes, damaged proteins do not accumulate in the cell. If "normal" mRNA were to be processed during heat shock, one can imagine that hsp's somehow stabilize the "spliceosome complex" that carries out the processing of precursor mRNA, thus allowing processing to occur at high temperatures. How hsp's do this is still not known. Other NIGMS grantees have suggested that hsp's have a structural role and may interact with other "normal" proteins, helping them to function under adverse conditions. Certainly this intriguing biologic phenomenon has many practical consequences, both positive, such as protecting cells against the damaging effects of fever, and negative, such as protecting tumor cells against heat therapy.



Joye F. Jones, Ph.D.

References

1. Yost, H.J. and S. Lindquist. (1986). RNA splicing is interrupted by heat shock and is rescued by heat shock proteins. Cell 45, 185-193.

2. Yost, H.J. and S. Lindquist. (1988). Translation of unspliced transcripts after heat shock. Science 242, 1544-1548.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 22, 1989

From Program Administrator, Genetics Program, NIGMS

Subject Research Highlight: "What Do Heat Shock Proteins Do?" R01 GM28988 (Walker, G.), Massachusetts Institute of Technology, and R01 GM36278 (Gross, C.), University of Wisconsin

To Director, NIGMS
Through, Director, Genetics Program, NIGMS 9-22-89

Cells exposed to higher-than-normal temperatures shut down a number of cellular processes including most gene transcription and protein production. Survival at higher temperatures depends on heat shock proteins (hsp's) whose functions are unknown. One set of genes, those encoding hsp's, are unusual because they are activated by elevated temperature and are transcribed at a high rate; their mRNA's are rapidly translated into protein during heat shock. The heat shock response has been found in every organism investigated, from archebacteria through eubacteria to eukaryotes including single celled and multi-celled organisms, plants and animals. Heat shock genes are highly conserved and the proteins they encode are very similar among organisms. Protection from heat damage can be "acquired" by shifting cells to higher temperatures in stages. For example, if bacteria growing at 25°C are incubated first at 30°C, they can withstand a temperature of 42°C that would be lethal, or at least severely damaging, if they were moved directly from 25°C to 42°C.

Heat shock proteins are required for cells to survive heat shock and other environmental stresses, but how they perform their protective functions is still a mystery. Evidence is accumulating that, in addition to their protective role, hsp's may be important for normal cellular function. An earlier Research Highlight (June 26, 1986) described work by Dr. Richard Voellmy who suggested that hsp's may somehow be related to normal degradation of damaged proteins. More recently, Dr. Carol Gross confirmed that hsp's are important in normal proteolysis and Dr. Graham Walker has evidence that they may also have a critical role in normal bacterial cell division.

Both scientists study a common bacterium, Escherichia coli, which has at least 20 heat shock genes. These genes can be activated by exposing cells to a large variety of environmental stresses including ethanol, DNA damaging agents, abnormal proteins and oxidizing agents, in addition to heat. Dr. Gross induced the heat shock response using a drug called puromycin that generates damaged proteins by terminating protein production prematurely. These damaged proteins are normally degraded by bacteria so that they do not accumulate and interfere with normal cell processes. By measuring the rate of degradation of these damaged proteins in bacteria with mutations in their heat shock genes, Dr. Gross was able to show that heat shock genes are necessary for normal proteolysis. Bacteria with a mutation in any one of five heat shock genes degraded damaged protein much more slowly than wild type bacteria with normal heat shock genes. Shifting the bacteria to higher temperatures increased the rate of proteolysis in both wild type and mutant cells, but, as expected, the mutants were still not as competent as wild type cells. This "negative" type of experiment, in which the lack of a protein interfered with a cell function, suggested strongly that hsp's are directly involved in proteolysis.

A more positive test for the role of these proteins in proteolysis would be to show that a higher concentration of hsp's enhanced protein degradation. Dr. Gross forced wild type cells to overproduce hsp's by a clever trick of genetic engineering. Heat shock genes in E. coli are transcribed by an RNA polymerase containing a sigma factor called sigma-32. (A sigma factor is one of the protein subunits of the polymerase that transcribes DNA, making the mRNA that is subsequently translated into protein; see Research Highlight of April 4, 1989, "To change your sigma, change your genes.") Sigma-32 is produced at very low levels unless bacteria are heat shocked. To increase the amount of sigma-32 in the cells, the gene for sigma-32 was placed on a plasmid under the control of a promoter regulated by a drug called IPTG. IPTG "turned on" the sigma-32 gene in bacteria containing these plasmids; sigma-32 protein then induced transcription of heat shock genes at a very high rate. These cells degraded proteins damaged by puromycin much faster than controls; as the temperature went up, so did the rate of proteolysis. These results confirm that hsp's are directly involved in proteolysis; furthermore, they suggest that the rate limiting step in proteolysis may be the amount of hsp in a cell.

Logically, hsp's could be proteases, especially since one heat shock gene, called lon146, is known to encode a protease. This is thought to be unlikely for several reasons. First, during heat shock, hsp's represent approximately 3% of the total cellular protein, which is a lot of protein for an enzyme. Secondly, hsp's are required for the replication of bacteriophages--in fact, they were first identified in mutants unable to support phage replication--and have been specifically implicated in the assembly of the phage head and tail; they are also required for disaggregation of the protein/DNA complexes involved in phage replication. None of these phage assembly functions are compatible with a protease. Finally, NIGMS grantee Graham Walker has shown that DnaK, the product of one heat shock gene, is required for normal bacterial cell division.

Dr. Walker found that E. coli cells completely lacking a functional dnaK gene are unable to grow at low or high temperatures and grow only slowly at an intermediate (30°C) temperature. Such cells grow as long filaments that are unable to divide normally. Bacteria divide by building a septum across the cell which divides it into two separate compartments that subsequently separate. The product of a gene called ftsZ, which is not a heat shock gene, induces septation. Cells without a functional dnaK gene can divide fairly normally if they have been engineered to overproduce FtsZ protein, although they grow more slowly than wild type cells. Thus, although cells that do not produce DnaK hsp can be forced to divide by laboratory tricks, they normally do not, indicating that this heat shock protein is required for making a cell wall.

Dr. Gross suggests that hsp's are structural proteins and Dr. Walker calls them chaperons. In either case, the scientists suggest that hsp's interact in key ways with other proteins that have a variety of different functions. The model may be pictured as scaffolding, in which hsp's provide a supporting structure enabling other proteins, particularly enzymes, to work efficiently. The scaffold may also be protective. Another NIGMS grantee, Susan Lindquist (see companion Research Highlight of September 22, 1989), has shown that hsp's are necessary for RNA splicing at elevated temperatures, and suggests that somehow hsp's protect the "spliceosome complex" from heat damage. Heat shock proteins are also important in development, although their role is not known. Now it is clear they are necessary for other cellular functions as well. They may represent a case of proteins needed for ordinary cellular housekeeping being conscripted to play a protective role when a cell is stressed.

Joyce F. Jones, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 4, 1987

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "per Jumps into the Gap" R01 GM 32329 (Young, M.), The Rockefeller University

To Director, NIGMS
Through: Director, Genetics Program, NIGMS GVB 9.9.87
Deputy Director, Genetics Program, NIGMS 9/9/87

Biological rhythms are universal, running the gamut from seasonal variations through daily (circadian) physiologic changes to very rapid (ultradian) oscillations. Researchers have studied biological rhythms in an enormous variety of organisms, including algae, sea slugs, fungi, and mice, but the most studied genetic system is the per (for period) locus of the fruit fly, Drosophila melanogaster. The product(s) of this locus is(are) involved in the control of circadian and ultradian rhythms and scientists suspect that this locus is a "master" clock gene. Two groups of scientists supported by NIGMS have isolated per and deduced that the gene product is a proteoglycan, but neither group knew what the per gene product actually does (see Research Highlight of July 16, 1986, "Clocks and Genes"). Now, in an exciting new development, Dr. Michael Young and colleagues at the Rockefeller University have evidence that per controls intercellular communications (1).

Scientists in Dr. Young's laboratory used radioactive RNA complementary to the mRNA of the per gene and fluorescent-tagged antibodies against the proteoglycan gene product to determine where the per gene is transcribed and translated during Drosophila development. While they confirmed earlier findings that per is expressed in neural tissues in adult flies, they also found that per is expressed preferentially in the salivary glands of embryos and that gene expression parallels the ontogeny of salivary gland development. This finding was not entirely unexpected since several years ago other scientists had reported the existance of circadian rhythms in isolated salivary gland cells (2). Because salivary gland cells have gap junctions, proteoglycans are known to modulate the function of gap junctions in hepatic cell cultures, and the product of per is a proteoglycan, Dr. Young reasoned that per may affect gap junctions. The cells of the salivary gland offered an excellent system to test his hypothesis.

Adjacent cells in a tissue may have impermeable junctions between them, junctions that are so tight that nothing can pass from one cell to another. Alternatively, cells may have tiny intercellular channels through which small

molecules (less than 1200 daltons) can pass; these channels are called gap junctions. A gap junction can be envisioned as a pipe connecting two tanks, a pipe that is about 30 angstroms long and 20 angstroms in diameter. While material within the cytoplasm can pass through these junctions, material from outside the cell cannot. Gap junctions are the conduits of chemical and electrical information between adjacent cells. (See Research Highlight of August 27, 1987, "Role of Gap Junctions in Developmental Patterning in Hydra".) Their function can be studied by injecting a small fluorescent dye into a cell and monitoring its appearance in adjacent cells and in the extracellular fluid. Additionally, gap junctions provide an area of low electrical conductance, so their function can also be monitored by passing an electrical current through one cell and measuring the voltage changes in adjacent cells. (This electrical conductivity of gap junctions accounts for the synchronous contractions of heart muscles.)

Dr. Young and colleagues studied gap junction function in salivary gland cells from wild type larvae and larvae with mutations in *per*. per^{0} mutants are arrhythmic whereas per^{s} mutants have shorter-than-normal rhythms. Cells from each type of fly were injected with a fluorescent dye; the dye moved more rapidly into adjacent cells in per^{s} mutants than in wild type cells, but did not move out of the injected cell in per^{0} mutants. Likewise, cells from per^{s} mutants conducted electricity much better than did those from wild type flies, whereas cells from per^{0} mutants conducted electricity more poorly. Thus, mutations in *per* dramatically affected the function of gap junctions, and hence the ability of cells to communicate.

The scientists do not know if *per* acts in nervous tissue the same way it does in salivary glands. Earlier work has shown that *per* must be functional in the brain for normal circadian rhythms in adults, and in the thorax for the normal ultradian rhythm of the male courtship song. Certainly the idea that *per* controls gap junctions is an attractive one for understanding how cells could behave rhythmically. Dr. Young suggests that the *per* gene product is not itself a central oscillator but rather functions as a "coupler", synchronizing various clock components in response to signals from some master control. This is because the major RNA transcript of *per* does not oscillate, but is present continuously. There are several other clock genes under study, most notably the *clock* and *andante* loci studied by NIGMS grantee Jeffrey Hall (PO1 GM 33205). At present, very little is known about these loci. One of them could be the master gene.

Joyce F. Jones, Ph.D.

References

1. Bargiello, T.A., L. Saez, M.K. Baylies, G. Gasic, M.W. Young, and D.C. Spray. 1987. The *Drosophila* clock gene *per* affects intercellular junctional communication. Nature 328: 686-691.

2. Weitzel, G. and L. Rensing. 1981. Evidence for cellular circadian rhythms in isolated dye-labelled salivary glands of wild type and an arrhythmic mutant of *Drosophila melanogaster*. J. Comp. Physiol. 143:229-235.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 19, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Cactus, Bugs and the Evolution of Drosophila"
R01 GM 34820-03 (Fogleman, J.), University of Denver

To Director, NIGMS
Through: Director, Genetics Program

Natural selection operates in complex environments making it extraordinarily difficult to define the characteristics that give an individual or species a fitness advantage over a competitor, and to understand the pressures that lead to the emergence of new species. An interesting system for studying this problem is the natural ecology of cactophilic species of Drosophila which share overlapping ranges in the Sonoran Desert, but which do not live in the same species of cactus. NIGMS grantee James Fogleman has been examining the relationship of cactus, Drosophila, and the microorganisms that cause cactus rot in which flies feed and breed.

The Sonoran Desert covers parts of the American Southwest, Baja California and northern Mexico. The four predominant cactophilic species of Drosophila, D. nigrospiracula, D. mettleri, D. pachea, and D. mojavensis live in the same area but occupy different niches by colonizing different species of cactus. The preference of the various Drosophila for specific cacti depends on the attractiveness of the volatile organics produced in rotting cactus and the ability of the flies to survive the potentially toxic compounds within the rot. D. pachea live in only a single kind of cactus, senita cactus. In Baja, D. nigrospiracula live preferentially in cardon cactus; this cactus is rare on the mainland so the flies live in saguaro cactus which is absent in Baja. D. mettleri preferentially live in the soil around these same two cacti, and consequently also shift hosts between Baja and the mainland; they are, however, "opportunistic generalists" that successfully adapt to a variety of cactus species and can live in tissues as well as soil. D. mojavensis were originally thought to prefer agria cactus both on the mainland and in Baja, even though agria were relatively rare on the mainland. Dr. Fogleman showed that D. mojavensis live equally well in the more numerous organ pipe cacti on the mainland.

Flies are attracted to cacti by volatile substances given off from rotting tissues. Dr. Fogleman's experimental system was to introduce injuries into the cactus with liquid nitrogen and study the microbial colonization of the rot, correlating it with the appearance of flies in the rot. In both natural and experimental systems, flies captured from the rot and larvae developing in the rot are 95% of a single species, the one that "prefers" that kind of cactus. Yeast have long been considered to be the microorganisms responsible for producing the volatile substances in the rot,

making cactus wine, as it were. Dr. Fogleman introduced pairs of injuries into organ pipe cactus, home to D. mojavensis, covering one with fine-mesh net and leaving the other open. Bacteria, mostly facultative anaerobes, rapidly colonized both injury sites;
yeast appeared later than bacteria and could be detected in open rots much sooner than in covered ones. Insects, including Drosophila, carried yeasts to the site of injury, and yeast in the covered sites probably arrived via cactus-dwelling mites that could crawl through the mesh. Bacteria, not yeast, seem to be the important organisms for producing the volatiles that attract flies to the rot. The first flies were detected at the site of injury after one day, and all uncovered rots had feeding flies and fly larvae by day 5. The maximum density of yeast in uncovered rots occurred 11 days after injury, well after flies were established. Further, the concentration of volatile attractants that the scientists identified in the rots were high enough by day 4 to attract flies, a time when the level of yeast colonization was still very low. The model is more one of cactus yogurt than cactus wine. Although bacteria generate the volatile compounds, the actual compounds produced depend more on the chemicals in the cactus tissues than on the species of bacteria. Dr. Fogleman found that the volatile attractants were mixtures of organic compounds rather than specific individual chemicals and that Drosophila species showed definite preferences.

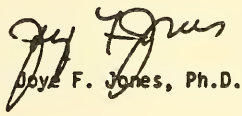
Once arriving at a particular cactus, the ability of a fly to feed and breed depends on the nutrients and toxic products in rotting tissue. The cacti that are home to desert-dwelling Drosophila belong to two biologic groups with different stem chemistries. D. pachea, D. nigrospiracula, and D. mettleri live and breed in one of these groups which includes saguaro, cardon, and senita cacti. These cacti contains alkaloids that are toxic to the other cactophilic drosophilids. D. mettleri are unusual in that their eggs and larvae develop in soil at the base of the plants that is soaked with cactus juice which has dripped off the stems and out of injuries. Evaporation causes the concentration of alkaloids to be up to 45 times higher than that found in rots on the plant, effectively excluding potential competitors that could survive in rots. Dr. Fogleman showed that high alkaloid concentrations were toxic to both adults and larvae of D. nigrospiracula, toxic to adults but only slightly toxic to larvae of D. mojavensis, and not toxic to adults but toxic only at the highest concentrations to larvae of D. mettleri. D. mettleri have a decided competitive advantage over the other species sharing the same range since they can survive in niches that kill competitors. Even though D. mojavensis would seem to be a potential competitor because their larvae can survive the high alkaloid concentrations, the alkaloid toxicity for adults prevents eggs from being laid; indeed, the levels of alkaloids found in the rots are toxic to D. mojavensis adults. Since D. mettleri breed in soil, D. pachea and D. nigrospiracula would appear to be major competitors for the same environment, the rots in the stems of these three cacti. However, D. pachea have an absolute requirement for a rare sterol found only in senita cactus, thus leaving cardon and saguaro cactus to D. nigrospiracula.

D. mojavensis preferentially live and breed in agria and organ pipe cactus, which belong to the second cactus subtribe. D. pachea cannot live in these two cacti because its essential sterol is absent, and D. mettleri preferentially breed in soil where they have no competition. Dr. Fogleman's research demonstrated that agria and organ pipe cacti contain a number of fatty acids and sterol diols which are toxic to D. nigrospiracula but not to D. mojavensis, thus eliminating the remaining competitor of D. mojavensis.

Speciation is the result of many different pressures, including geologic changes and climatic shifts. The well-studied differences among the Hawaiian Drosophila species are considered to be a result of geographic isolation resulting from the volcanic formation of the islands and subsequent colonization of each island by a small number of founder individuals.

Speciation of the cactophilic Drosophila, arising probably from a climatic change which altered the natural flora, reflects the differential capacities of the various flies to respond to different chemicals produced in rotting cactus and to survive the chemical milieu of the rots. Studies like those of Dr. Fogleman are important in trying to understand the pressure that lead to reproductive isolation among organisms that share a geographical range.

Human variation may be a result of relatively small environmental differences occurring in different areas during man's evolution. Human differences have not led to reproductive isolation, but the same kinds of pressures that lead to speciation can select for human differences. Understanding these pressures in model systems is useful for understanding the genesis of human variation.



Joyce F. Jones, Ph.D.

References

Meyer, J.M. and J.C. Fogleman. 1987. Significance of saguaro cactus alkaloids in ecology of Drosophila mettleri, a soil-breeding cactophilic drosophilid. J. Chem. Ecol. 13, 2069-2081.

Fogleman, J.C. and L. Armstrong. 1988 Ecological aspects of cactus triterpene glycosides. I. Their effect on fitness components of Drosophila mojavensis. J. Chem. Ecol. In press.

Fogleman, J.C. and W.B. Heed. 1989. Columnar cacti and desert Drosophila: The chemistry of host plant specificity. In Interactions among Plants and Animals in the Western Deserts, J.D. Schmidt, ed. (Univ. New Mexico Press) In press.

D. MEDICAL APPLICATIONS OF BASIC RESEARCH

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date March 20, 1989

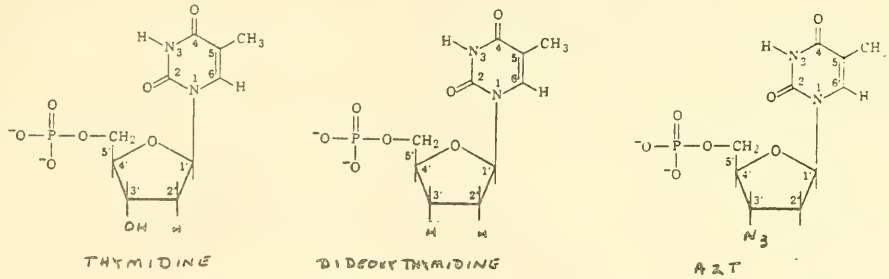
From Program Administrator, Genetics Program, NIGMS

Subject HIGHLIGHT: "The Basis of AZT Action, mtDNA Replication, and Potential Drugs Against AIDS. GM 22333-21 (Simpson, M.), State University of New York at Stony Brook

To Director, NIGMS
Through: Director, Genetics Program, NIGMS

AZT (3'-azido-3'-deoxythymidine) is a potent inhibitor of HIV-1 replication in vitro and is widely used as a therapy for AIDS. The usefulness of the drug is limited by its toxic side effects, particularly that of bone marrow depression. Recent research by Dr. Melvin Simpson suggests that the toxicity of AZT may be due to its effects on mitochondrial DNA replication. Further, his research identifies several related chemicals which have potential as drugs for the treatment of AIDS.

The structures of thymidine, dideoxythymidine (ddTTP), and AZT are shown below:



Replication of DNA proceeds by addition of nucleotides to the 3' carbon. Some analogs of the nucleotides, such as ddTTP and AZT, block DNA replication because they lack an attachment site for the next 3',5'-phosphodiester bond. It follows that the better a DNA polymerase utilizes such an analog as a substrate, the greater will be the inhibition of DNA replication. The effectiveness of AZT against HIV-1 stems from the propensity of the HIV-1 reverse transcriptase to utilize AZT as a substrate. Nuclear DNA replication is resistant to AZT because AZT is not readily utilized by DNA polymerase α , the primary nuclear DNA polymerase.

In the late 1960's, Dr. Simpson isolated and characterized DNA polymerase γ, which later was shown to be responsible for mitochondrial DNA replication. This early characterization showed that DNA polymerase γ avidly utilizes dideoxynucleotides. Dr. Simpson wondered if the toxic effects of AZT might result from its inhibition of mtDNA replication in bone marrow cells. Even a small inhibition of mtDNA replication could have serious consequences for the viability of cells. Dr. Simpson's research program was, therefore, designed to address the following two questions:

Does AZT inhibit mtDNA replication?

How do other chain terminators, especially those known to have anti-HIV activity, affect mtDNA replication?

Mitochondrial DNA replication is typically determined by measuring the amount of a labeled DNA presursor, such as tritiated thymidine or tritiated dATP, incorporated into isolated mitochondria. Dr. Simpson assayed mtDNA replication in the presence of AZT or one of ten other chain terminators. Each chain terminator was tested twice, once in the presence of tritiated thymidine and once in the presence of tritiated dATP, to assure that the labeled precursors did not compete with the chain terminators and obscure their effects on mtDNA replication. In fact, the results of experiments in which tritiated thymidine was used to measure mtDNA replication were the same as those in which tritiated dATP was used. This gave Dr. Simpson the confidence that his experiments measured real effects of AZT and other analogs on mtDNA replication.

The results of Dr. Simpson's experiments were striking, as shown below.

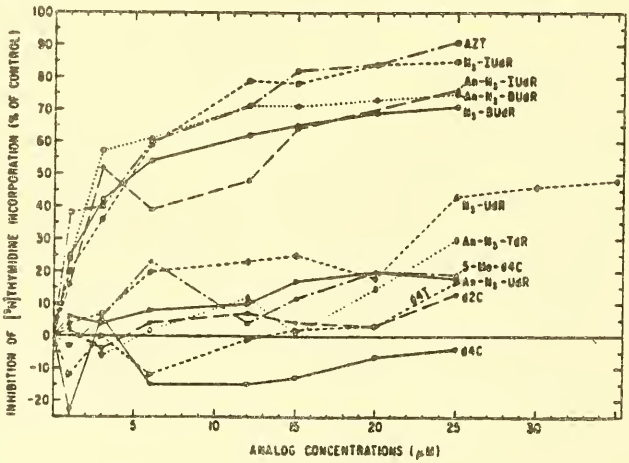


Fig. 2. Effects of nucleosice analogs on [^{3}H]thymidine incorporation into mtDNA by isolated rat mitochondria. Control values were about 500 cpm.

AZT (at the top of the graph) is a potent inhibitor of mtDNA replication. The eleven chain terminators tested fell into two groups, one much stronger than the other in inhibiting mtDNA replication. The chemicals in the low group have potential as AIDS drugs because of the possibility that they would be less toxic.

Dr. Simpson is quick to point out that although these results are suggestive, many more tests are necessary to answer a variety of questions. For example, this work does not explain the highly selective toxicity of AZT against bone marrow. Dr. Simpson is now working with Friend mouse leukemia cells, which can be induced to produce hemoglobin. Preliminary experiments show that uninduced Friend cells are resistant to high concentrations of AZT but that they are very sensitive after induction.

One of the chain terminators shown by Dr. Simpson to be a weak inhibitor of mtDNA replication was tested as a therapeutic agent against AIDS in 1988. As Dr. Simpson would have predicted, it evoked no signs of bone marrow toxicity; unfortunately, it did have other serious side effects. Two other weak inhibitors are now being studied in clinical trials.

This work provides another excellent example of how basic, untargeted research can have important practical consequences. Over the past 21 years, Dr. Simpson's work has helped scientists understand the molecular organization and function of mitochondrial DNA. It now has the potential for important contributions to the development of AIDS therapies.

Irene Anne Eckstrand

Irene A. Eckstrand, Ph.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 19, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Chromosomal Fragile Sites: From Drosophila to Humans"
R01 GM 19179 (Laird, C.), University of Washington

To Director, NIGMS
Through: Director, Genetics Program 9-20-88

Fragile sites in chromosomes of humans and other organisms are sites that break readily when the cells are placed in culture. Cytologically, the sites often appear as constrictions or gaps in the chromosome. The phenotypic effects of fragile sites vary from no observable effects in Drosophila to being directly correlated with mental retardation and tumorigenesis in humans. One of the best-studied fragile sites in humans is Xq27, which is located on the X chromosome. Xq27, referred to as fragile X, is associated with the most common form of inherited mental retardation. The inheritance of the fragile X has long been perplexing. Fragile X-associated mental retardation occurs primarily in males, but in order for the syndrome to be expressed, the affected chromosome must have been inherited from the mother. Furthermore, the defect displays variable penetrance, which means that the severity of the disorder varies in different individuals.

Dr. Charles Laird has been studying intercalary chromatin, the equivalent of human fragile sites, in Drosophila salivary gland polytene chromosomes. Polytene chromosomes are giant chromosomes that form as a result of multiple rounds of DNA replication without cell division. Dr. Laird's work began as a test of his theory that delayed replication was responsible for intercalary chromatin in polytene chromosomes (1, 2). He used molecular techniques to measure the time of replication of intercalary chromatin and found that it was delayed compared to DNA in other chromosomes. Additional experiments with human/rodent hybrids, in which replicating human chromosomes were placed in a cell undergoing chromosome condensation, indicated that intercalary chromatin formed at sites that had not completed DNA synthesis when the chromosome was forced to begin condensation.

Having demonstrated that in two other organisms gaps can be correlated with delayed replication of DNA and premature condensation of chromatin, Dr. Laird searched the literature for similar data in the formation of human fragile sites. He found that the methods used to increase the frequency of fragile sites in cultured human cells involved shortening the G2 phase (the stage between DNA synthesis and mitosis) of the cell cycle. The effect of a shortened G2 phase is to decrease the time between the completion of DNA replication and chromosome condensation, analogous to the cell fusion experiment. Numerous agents that inhibit DNA synthesis in human cells have been reported to produce fragile sites in human chromosomes. Based on his

own results and the data in the literature about human fragile X, Dr. Laird proposed that fragile sites in humans result when late-replicating DNA is not allowed sufficient time to condense properly. The regions that are incompletely condensed form the fragile site. The rest of the chromosome, however, replicates on time and condenses normally.

If late replication resulting in inappropriate condensation is responsible for the formation of fragile sites, can other observed properties of fragile X be explained by this model as well? Dr. Laird believes so, although some in the medical genetics community are still reluctant to accept his interpretation. He has proposed the following explanations for several fragile X properties:

> The cytological appearance of fragile sites (usually light staining or the presence of a gap) is the result of altered condensation.
>
> The fragility of fragile X sites is due to breakage of the incompletely condensed DNA where there are fewer strands per unit area when force is applied.
>
> The recombinogenicity of fragile sites results from the presence of single-stranded regions at the replication fork remaining in the fragile site, where nonhomologous pairing between chromosomes could occur.

Since the known characteristics of fragile X sites could be explained by Dr. Laird's model, his next step was to suggest why the DNA at fragile X sites replicates late. Two pieces of data had to be taken into account. First, not all fragile sites are suppressed or induced by the same agents. Thus, there must be multiple DNA alterations that lead to fragile site formation. Second, human/hamster hybrid experiments have demonstrated that the presence of other human chromosomes is not necessary for the maintenance of the fragile X state. Therefore, the alterations are not caused by factors from other chromosomes.

Dr. Laird has proposed a model in which two alterations of the X chromosome cause fragile X formation. The first alteration is a mutation in the DNA, resulting in late replication of the DNA at that site. The second alteration is methylation of the DNA. Methylation is a particularly attractive explanation for fragile site formation, since agents that demethylate DNA also eliminate fragile sites.

Dr. Laird proposes that when a mutation in the DNA results in late replication of that DNA during oogenesis, the reactivation of the site is interrupted. Reactivation of an inactive X chromosome occurs during oogenesis in the female when the methyl groups on the DNA are removed. Studies have shown that it is the methylation of the DNA that suppresses gene expression in the inactive X chromosome. When the DNA at the site of the mutation replicates late, the demethylation of the rest of the X chromosome has already occurred and the fragile site is never demethylated. Then the fully methylated DNA containing the mutation replicates even later, since methylated DNA often replicates late. At this stage the cytologically visible fragile site would be apparent

when the late replicating DNA is forced to condense before it is completely replicated. The methylation pattern of the fragile site is then inherited by offspring, a process called imprinting. Details of the model are given in Figure 1.

The attractive feature of this model is that the failure of part of the X chromosome to reactivate during oogenesis explains why the fragile X must be passed through a female in order to be expressed. The imprinted fragile X represents the case with the most severe phenotype and highest induction rate. However, there are intermediate levels of expression, and Dr. Laird expects to find that these represent cases where the DNA is not replicated as late in the cell cycle.

Dr. Laird's model has met with a good deal of resistance (3, 4). One argument has been that many fragile sites are known to reside in DNA that is not replicated late in the S phase. The model, however, is not dependent upon the normal site residing in late-replicating DNA, but rather upon a mutation at the site that causes late replication of the DNA. Numerous other points have been used to argue against the model, but in most instances, Dr. Laird's model has been consistent with those points. Dr. Laird plans experiments on the timing of replication in human cells that will be very important in establishing the validity of the model.

Dr. Laird's work is an excellent example of how basic research in a lower organism can lead to insights into human disorders. A molecular understanding of chromosome reactivation may result in the development of techniques to restore normal replication and transcription in active DNA. Such developments could have significant clinical applications.

Jane L. Peterson, Ph.D.

References

1. Lamb, M. M. and Laird, C. D. (1987). Three euchromatic DNA sequences under-replicated in polytene chromosomes of Drosophila are localized in constrictions and ectopic fibers. Chromosoma 95, 227-235.

2. Laird, C., Jaffe, E., Karpen, G., Lamb, M., and Nelson, R. (1987). Fragile sites in human chromosomes as regions of late-replicating DNA. Trends in Genetics 3, 274-281.

3. Hecht, F. (1988). Enigmatic fragile sites on human chromosomes. Trends in Genetics 4, 121.

4. Laird, C., Jaffe, E., Karpen, G., Lamb, M., and Nelson, R. (1988). Reply from Charles Laird et al. Trends in Genetics 4, 122.

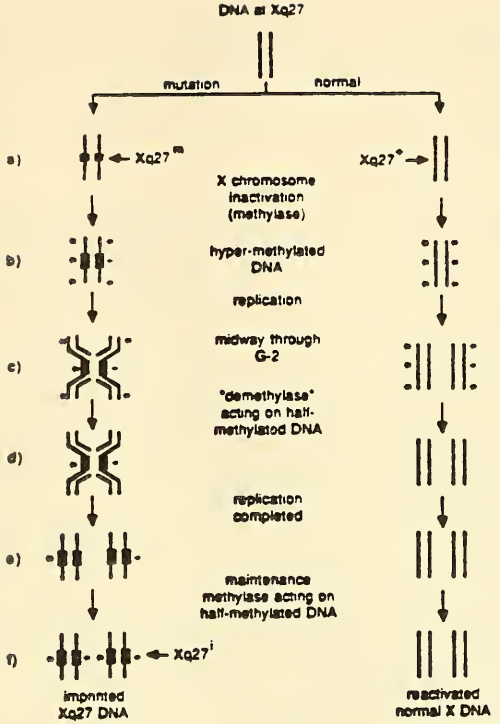


Fig. 1 *Proposed mechanism by which late replication blocks the process of X chromosome reactivation prior to oogenesis. (a) A mutation (called 'pre-mutation' in Refs 65, 66) changes normal DNA at Xq27 (Xq27$^+$) to a mutated allele Xq27^m, which replicates later in the cell cycle. (b) As part of the normal process of X chromosome inactivation in female cells[64], the inactive X chromosome becomes methylated. (c) In preparation for the reactivation that occurs prior to oogenesis[71,72], one cell division occurs without maintenance methylase activity. (d) A 'demethylase'[60], or a DNA glycosylase that replaces 5-methylcytosine by cytosine[60], removes methyl groups from half-methylated DNA. (e) DNA synthesis is finally complete at Xq27, but it is too late for reactivation of that region because half-methylated DNA was not available during the time that 'demethylase' was present. (f) In a subsequent cell cycle, maintenance methylase activity is again present, restoring full methylation to the DNA at fragile site Xq27. At this time the fragile X chromosome is considered to be 'imprinted'[19]. The imprinted allele, Xq27', is even later-replicating, and it is now more easily detectable as a chromosome gap under conditions of fragile site induction. The imprinted allele Xq27' is then stably inherited through both males and females[19]. For this illustration, the mutated fragile X chromosome is indicated as imprinted upon completion of the cycle of inactivation/reactivation. It is not known, however, whether or not the entire cycle must be completed before imprinting is observable cytologically.*

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date October 30, 1989

From Director, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT - "Drosophila May Provide Clues to the Metastatic Process" R01 GM33959 (Shearn, Allen) Johns Hopkins University

To Director, NIGMS

During early embryonic development of Drosophila, the somatic cells segregate into larval and imaginal lineages. When the larva hatches from the egg, the larval cells grow in size while the imaginal cells act as stem cells, undergoing many cycles of cell division. Although the imaginal cells remain embryonic in appearance, they become programmed for future differentiation. Then, during metamorphosis, when most larval tissue degenerates, the thousands of imaginal cells, grouped together in structures called imaginal discs, give rise to a variety of adult structures.

Dr. Allen Shearn and members of his laboratory have been interested in identifying genes which affect imaginal disc development. His strategy for identifying these genes is to isolate Drosophila with mutations that cause the occurrence of defective discs or of abnormal structures derived from the discs. One of the genes he discovered in his mutant search was abnormal wing discs (awd), so named because of the gross defects in the wing imaginal discs of flies with this mutation; less severe effects were also seen in the structures that develop from the eye-antenna and leg discs. Based on the developmental consequences of a defective awd gene, Dr. Shearn inferred that awd is not required during embryogenesis or early larval development but is required during the second half of larval development, both in the wing disc and brain and also, somewhat later, in the eye-antenna and leg discs and in the ovary (1). From studies in which normal imaginal discs were grown in mutant larvae and mutant imaginal discs were grown in normal larvae, Dr. Shearn's group showed that the awd gene is cell-autonomous, which means that mutations in the gene affect only the cells in which awd is normally expressed. These studies also led to the suggestion that the defect in wing discs in awd mutants is due to abnormal cell death, although the abnormalities observed in the other tissues are probably due to other causes.

Dr. Shearn's group next isolated genomic DNA, which they showed corresponded to the awd gene (2). The gene encodes a 0.8 kilobase RNA transcript which shows the highest accumulation in normal larvae just prior to the stage at which the mutant awd larvae begin to express an abnormal phenotype. As expected, the RNA was found in all of the tissues which are affected by the awd mutation, as well as in at least one tissue (salivary gland) in which no defect was observed.

Once the awd gene was cloned, Dr. Shearn determined its DNA sequence, and as most investigators do with a new sequence, compared it to sequences in GenBank. To his surprise, he discovered that a human gene, nm23, which encodes a 17 kilodalton protein, is 78% identical to awd throughout the entire translated sequence. The nm23 gene was originally identified by investigators in the laboratory of Dr. Lance Liotta in the National Cancer Institute, who showed that it is differentially expressed in both rodent and human melanoma cell lines with varying metastatic potentials (3,4). Levels of nm23 are uniformly reduced in cell lines with high metastatic potential. Metastases from human breast carcinomas also contain reduced levels of nm23 RNA. The nm23 protein is localized in the nucleus and cytoplasm.

Dr. Shearn's laboratory and Dr. Liotta's laboratory collaborated to compare the awd and nm23 genes (5). They found that, although the translated sequences of the two genes are highly conserved, the flanking sequences have little similarity. As expected from the high degree of homology between the coding regions of the human and Drosophila genes, antibodies to an nm23 peptide recognize a 17 kilodalton protein in lysates of Drosophila embryos.

It is not uncommon these days to find highly homologous genes in very diverse species. What makes this serendipitous discovery exciting, however, is that the phenotype of wing imaginal discs in awd mutants parallels the phenotype of cancer cells in which nm23 expression is reduced. In both cases, there is altered cell morphology, aberrant differentiation, cell necrosis, and heterogeneity in the pattern of abnormal morphology and cell death. Furthermore, Dr. Shearn has recently found that the awd protein, like the nm23 protein, is localized in both the nucleus and the cytoplasm. Thus, the nm23/awd gene may contribute to the normal development of cells, and loss of expression of nm23/awd may lead to a disordered state, favoring aberrant development or cancer progression.

The main take-home messages of this story is that basic research often has unexpected payoffs. Dr. Shearn never dreamed, when he embarked on studies of Drosophila imaginal disc development, that he would stumble onto new insights into the metastatic process in cancer and that his discovery might ultimately result in new treatments for the disease.

Judith H. Greenberg, Ph.D.

References:

1. Dearolf, C.R., Hersperger, E., and Shearn, A. (1988) Developmental consequences of awd^{b3}, a cell-autonomous lethal mutation of Drosophila induced by hybrid dysgenesis. Develop. Biol. 129:159-168.

2. Dearolf, C.R., Tripoulas, N., Biggs, J., and Shearn, A. (1988) Molecular consequences of awd^{b3}, a cell-autonomous lethal mutation of Drosophila induced by hybrid dysgenesis. Develop. Biol. 129:169-178.

3. Steeg, P.S., Bevilacqua, G., Kopper, L., Thorgeirsson, U.P., Talmadge, J.E., Liotta, L. A., and Sobel, M.E. (1988) Evidence for a novel gene associated with low tumor metastatic potential. J. Natl. Cancer Inst. 80:200-204.

4. Bevilacqua, G., Sobel. M.E., Liotta, L.A., and Steeg, P.S. (1989) Association of low nm23 RNA levels in human primary infiltrating breast carcinomas with lymph node involvement and other histopathological indicators of high metastatic potential. Cancer Res. 49:5185-5190.

5. Rosengard, A.M., Krutzsch, H.C., Shearn, A., Biggs, J., Barker, E., Margulies, I.M.K., King, C.R., Liotta, L.A., and Steeg, P.S. (1989) Reduced nm23/awd protein in tumor metastasis and aberrant Drosophila development. Nature (in press).

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date November 29, 1988

From Program Administrator, Genetics Program, NIGMS

Subject Research Highlight: "New Mutations Cause An Old Disease" ROI GM-28931, Alan F. Scott, John Hopkins University, Baltimore, Maryland

To Director, NIGMS
Through: Director, Genetics Program 11-29-88

Hemophilia is one of the most infamous genetic disorders, having become particularly well known among the descendents of Queen Victoria, most notably, Victoria's great grandson, the son of the last Russian Czar. Hemophilia is a disorder of the blood clotting system in which a critical component, Factor VIII, is missing or non-functional. The disease is inherited as an X-linked recessive condition, commonly expressed in men but transmitted by women who themselves are rarely affected. Half of the daughters of carriers are also carriers, half of their sons have the disease, and their other children are normal. Most individuals with hemophilia have a familial history of the disease, but a significant number of cases represent newly arisen mutations. More than 200 such cases have been found at the Johns Hopkins Unversity, and most were caused either by deletions or point mutations. NIGMS grantee Alan F. Scott and his colleagues from Hopkins have recently described two unrelated patients whose hemophilia appears to have arisen by insertion of DNA into one of the exons, exon 14, in the gene which encodes Factor VIII. (Exons are those parts of genes containing the sequences for the mRNA which in turn will be translated into the protein gene product; exons are separated from each other by non-coding sequences called introns. See Research Highlight of March 30, 1988, "What's a nice intron like you doing in a gene like this?") In both patients, the inserted sequences belong to a family of repetitive DNA sequences called L1 sequences.

Mammalian genomes contain many copies of repetitive DNA sequences whose function is unknown. They are classified as either long or short, and in humans, the best studied long sequences are called L1. They are present in 10,000 to 100,000 copies scattered throughout the genome. A full-length L1 sequence is about 6 kilobases, but most are shorter, lacking DNA from the 5' end. L1's have an adenine-rich region at the 3' end and contain two open reading frames, one of which (ORF-2) encodes a polypeptide with high sequence homology to reverse transcriptase, an enzyme that catalyzes the production of DNA from an RNA template. Consequently, L1's are thought to represent a non-viral class of retrotransposons, that is, DNA sequences derived from mRNA (mRNA has a long 3' polyadenosine tail) which have inserted randomly into the genome.

The two new mutations described by the Hopkins group were intriguing because, in both cases, a truncated L1 had inserted into exon 14. Although each insertion was in a different site within the exon, each had inserted in the midst of a region rich in adenines. The researchers suggest the following hypothesis to account for this curious observation. In a germ-line cell, one (or a small number) of the full-length L1 sequences is transcribed making mRNA which is then translated into protein. One protein product is a reverse transcriptase (from ORF-2) which in turn "reverse transcribes" mRNA back into DNA (functionally, a cDNA). Some of these cDNA's would be full length, but many would be shortened, perhaps because the enzyme is inefficient. The poly-A "tail" of mRNA becomes a string of T's in the cDNA. This cDNA would then be able to insert into genomic DNA at a complementary site, specifically a region rich in A's, such as is found in exon 14 of the Factor VIII gene. The inserted cDNA would be copied onto the opposite strand, now ensuring this "rogue" DNA's place in the genome. Although this DNA could insert within any string of A's, it would have an effect only if the insertion occurred within a gene or its regulatory regions.

This is the first report of a human genetic disease caused by insertion of a large DNA sequence into a nonadjacent site (thus differing from tandem duplication) and involving, in all likelihood, an RNA intermediate. The frequency of this kind of mutation, called retrotransposition, is unknown, but of the 240 new mutations discovered by the Hopkins group, two, or almost 1%, were derived in this way. These observations suggest that this mode of mutation is fairly common. Scientists frequently discuss mutations caused by single base changes, deletions large and small, duplications, and even small insertions. Mutation by retrotransposition involves inserting a large piece of DNA into a gene, and certainly will make scientists consider another form of change when considering the underlying cause of a genetic disease.

Judy F. Jones

References:

Kazazian, H.H., Jr., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips, and S.E. Antonarakis. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332, 164-166.

Scott, A.F., B.J. Schmeckpeper, M. Abdelrazik, C.T. Comey, B. O'Hara, J.P. Rossiter, T. Cooley, P. Heath, K.D. Smith, and L. Margolet. 1987. Origin of the human L1 elements: Proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1, 113-125.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date September 19, 1988

From Program Administrator, Genetics Program, NIGMS

Subject RESEARCH HIGHLIGHT: "Fungi or Protozoa: Only the DNA Knows for Sure"
R01 GM 32964 (Sogin, M.) National Jewish Center for Immunology and Respiratory Medicine

To Director, NIGMS
Through: Director, Genetics Program, NIGMS

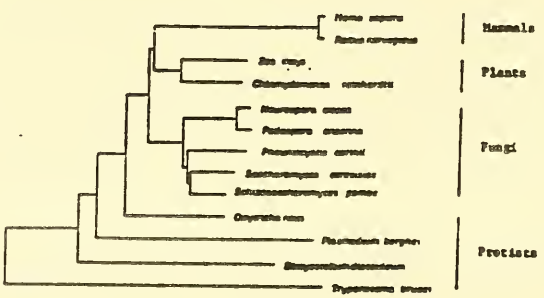
Pneumonia caused by the ubiquitous Pneumocystis carinii is the leading cause of death in patients with AIDS and contributes significantly to death in patients immunosuppressed for other reasons such as transplantation or cancer chemotherapy. The organisms can live in a number of different mammalian hosts, and most healthy people carry them in their lungs with no obvious problems. Colonization is so common that the organisms can be considered normal inhabitants of the lungs. Like many such organisms, P. carinii are opportunists, and when immunity is compromised, they cause serious and often fatal interstitial cell pneumonia.

The biologic classification of P. carinii has been uncertain for a long time. The life cycle is unknown, although several different morphological stages have been described. So far, the organisms cannot be cultured in vitro, making it more difficult to study their biology. Organisms are transmitted primarily, if not exclusively, by air, but the infectious stage is not known and person-to-person transmission has not been demonstrated. Since organisms can be visualized in tissues by stains usually used for fungi and their "cell walls" are fungal-like, some biologists classify P. carinii as fungi. However, since they respond to drugs used primarily to treat protozoal infections and some of their organelles are like those of protozoa, P. carinii have more recently been considered to be protozoa. These particular anti-protozoan drugs (pentamidine or the combination of trimethoprim and sulfamethoxazole) are now known to be effective against fungi, so the classification may be in error. The importance of accurate classification rests in the ability to design rational drugs that are more effective and less toxic, and to define the life cycle, particularly identifying the infectious stage of the organism and its source. If P. carinii are protozoa, their life cycle would be expected to be like those of other protozoa and may involve a reservoir or intermediate host. If they are fungi, their life cycle may be similar to that of fungal pathogens such as Histoplasma capsulatum in which the infectious form and the form that causes disease (histoplasmosis) are not the same.

Using techniques developed primarily to understand evolutionary relationships among lower eukaryotes, NIGMS grantee Mitchell Sogin has recently found that P. carinii are more closely related to fungi than to protozoa (1). Geneticists interested in understanding evolution have been using DNA sequences as tools for building family trees among highly divergent species.

Dr. Sogin has long been interested in the genetic relationships among eukaryotes, particularly the single cell organisms whose evolutionary past leaves no fossil record and whose structures or ontogeny give few clues about biologic relationships. Several years ago, Dr. Sogin and colleagues (in particular, a group led by C. R. Woese and G. E. Fox) developed a method of comparing genomes by analyzing the complete sequences of small subunit ribosomal RNAs (16S-like rRNAs). These rRNAs have regions that are highly conserved, regions that are semi-conserved, and regions that are non-conserved. The highly conserved regions are useful for developing probes to find rRNA genes in a genomic library. The semi-conserved regions are useful for building phylogenetic trees between distantly related organisms. The non-conserved regions can be used for resolving close phylogenetic relationships. The accumulation of differences in DNA sequences of the rRNA genes is a function of the time since the various progenitor organisms separated to begin their individual evolutionary journeys (2).

Dr. Sogin and his group of collaborators created DNA probes from conserved regions of rRNA genes from Escherichia coli which would identify the rRNA genes in a genomic library from P. carinii isolated from the lungs of immunosuppressed rats. (When housed in conventional rather than germ-free environments, these rats have large numbers of P. carinii in their lungs.) The DNA fragments identified by the bacterial probes were cloned and the genes sequenced. The sequences were aligned with sequences of rRNA genes from a large number of eukaryotes, including "higher" organisms such as rats and corn; various fungi including bread mold and baker's yeast; and protozoa including plasmodia (which cause malaria) and African trypanosomes. Alignments were initially made at the conserved regions; computer algorithms were used to align less conserved and non-conserved regions and to deal with gaps. Using the differences in sequences, including gaps, the scientists calculated the frequency of change at each nucleotide position and constructed the phylogenetic tree shown in the figure. This analysis places Pneumocystis carinii near the yeasts (Saccharomyces cerevisiae and Schizosaccharomyces pombe) and their fungal relatives (Neurospora crassa and Podospora anserina) but very far from the ciliated protozoan Oxytricha nova, the malarial protozoan Plasmodium berghei, and the kinetoplastid protozoan Trypanosoma brucei.



Phylogenetic tree inferred from 16S-like rRNA sequence similarities. The evolutionary distance between nodes of the tree is proportional to the length of the horizontal line.

This work gives us useful and critical information about a very serious human disease at a crucial time in medical history. Equally important, however, is the validation of the NIGMS policy of supporting basic research for its own sake. Dr. Sogin did not set out to study human disease per se; his goal was to understand a fundamental biological problem: evolutionary relationships of lower eukaryotes. Serendipitously, at a critical point in the AIDS epidemic, his work has provided insight into an organism that is causing death in thousands of people. These findings will enable researchers and clinicians to look at Pneumocystis pneumonia in a new way, think of new approaches to treatment, and perhaps eliminate it altogether. There can be no better payback.

Joye F. Jones, Ph.D.

References:

1. Edman, J.C., J.A. Kovacs, H. Masur, D.Y. Santi, H.J. Elwood, and M.L. Sogin. 1988. Ribosomal RNA sequence shows Pneumocystis carinii to be a member of the Fungi. Nature 334, 519-522.

2. Sogin, M.L. and J.H. Gunderson. 1987. Structural diversity of eukaryotic small subunit ribosomal RNAs. Ann. N.Y. Acad. Sci. 503, 125-139.

MINORITY ACCESS TO RESEARCH CAREERS (MARC) PROGRAM

OBJECTIVES

The Minority Access to Research Careers (MARC) Program was formally established in 1975, and is a special research training support program administered by NIGMS in collaboration with other institutes of the NIH. The primary goals of the MARC Program are to increase the number and capabilities of minority scientists engaged in biomedical research and to strengthen science curricula and research opportunities at minority institutions in order to prepare students for careers in biomedical research. To achieve these goals, the MARC Program uses both institutional research training grants and individual fellowships. These are: the MARC Faculty Fellowship, the MARC Visiting Scientist Award, the MARC Honors Undergraduate Research Training Grant, and the MARC Predoctoral Fellowship.

PROGRAM COMPONENTS

The MARC Faculty Fellowship, the first of the four MARC support mechanisms developed, provides opportunities for advanced research training for selected faculty members of 4-year colleges, universities, and health professional schools in which student enrollments are drawn substantially from ethnic minority groups (such as American Indians, Blacks, Hispanics, and Pacific Islanders). These institutions may nominate faculty members for MARC fellowships for a period of up to 3 years of advanced study and research training, either as candidates for the Ph.D. degree or as investigators obtaining postdoctoral research training in the biomedical sciences. Faculty Fellows are expected to return to the nominating institutions at the completion of their training.

MARC Faculty Fellows are selected on a competitive basis. Evaluation is based upon the applicant's qualifications and potential for research and training, as evidenced by academic records, reference reports, and publications, as well as by the research training proposal, the proposed training situation (i.e., the sponsor's and the institution's training facilities and staff), and other relevant information.

The second support mechanism, the MARC Visiting Scientist Award, provides financial support for outstanding scientist-teachers to serve as visiting scientists at 4-year colleges, universities, and health professional schools where student enrollments are drawn substantially from ethnic minority groups. The primary intent of this award is to help strengthen research and teaching programs in the biomedical sciences for the benefit of students and faculty at these institutions by allowing them to draw upon the special talents of expert scientists from other institutions. Reciprocal benefits should accrue to the Visiting Scientist through the added experience gained by his or her involvement in innovative science education and research development programs. The proposal must include arrangements for the Visiting Scientist to reside in the campus community and to participate fully in programs of teaching, development of research, and/or counseling, as outlined by the institution. Evidence of negotiations and of some tentative agreement between the applicant institution and the scientist-teacher nominated to serve as the MARC Visiting Scientist should be provided as part of the application. The individuals nominated should be recognized scholars and

leaders in the biomedical sciences. Proposals may request support for periods ranging from 1 academic quarter to a maximum of 1 year. Stipends are determined on an individual basis, according to the nominee's current salary or other possible source of stipend support for the proposed period in residence.

The MARC Honors Undergraduate Research Training Grant, the third mechanism of support, was initiated at the suggestion of Congress and Institute consultants and staff as a means of emphasizing the value and importance of providing biomedical research training at the undergraduate level in minority institutions. The objectives of the program are: to increase the number of well-prepared students who can compete successfully for entry into graduate programs leading to the Ph.D. degree in the biomedical sciences, to help develop a strong undergraduate science curriculum, and to stimulate an interest in undergraduate research as preparation for graduate study in the biomedical sciences.

Training support is offered to carefully selected undergraduate honors students at 4-year colleges, universities, and health professional schools in which student enrollments are drawn substantially from ethnic minority groups. Each institutional grant is awarded for a maximum period of 5 years. These honors programs for third and fourth-year students should be designed to improve significantly the research training capabilities of the minority institutions. Applications for support should provide information regarding proposed mechanisms to augment and improve the science curricula, strengthen the faculty, and improve laboratory facilities. Funds are available for research equipment and supplies essential to the program, stipends, tuition and fees, and limited travel costs for the trainees. Specific arrangements for special training at universities and laboratories other than the grantee institution should be described in the application.

The fourth component of the program, instituted in January of 1981, is the MARC Predoctoral Fellowship, a National Research Service Award that provides support for research training leading to the Ph.D. degree in the biomedical sciences for selected students who are graduates of the MARC Honors Undergraduate Research Training Program. It is expected that such training will be conducted in graduate-degree programs of the highest quality. Support is not available to individuals enrolled in medical or other professional schools, unless they are enrolled in a combined degree (M.D.-Ph.D.) program. Awards are conditional upon acceptance into a specific doctoral degree (Ph.D.) program in biomedical research.

MARC Predoctoral Fellows are selected on a highly competitive basis. A maximum of 5 years of support may be recommended, based on the merit of the application and evidence of satisfactory progress in the doctoral program in which a successful applicant is enrolled. The award provides an annual stipend to the student and funds to help defray the expenses of tuition, fees, and supplies.

ORGANIZATION AND STAFFING

The MARC Program is administered by its Director, Mr. Elward Bynum, and one professional staff member, Mrs. Dolores L. Lowery, as program administrator.

HIGHLIGHTS

Since the inception of the MARC Faculty Fellowship Program in 1972, approximately 450 applications for MARC fellowship support have been received and 220 individuals have received support under the program. These individuals include 124 predoctoral and 96 postdoctoral award recipients. Overall, 90 percent of those who have completed the program have returned either to the original home institution or to another minority institution.

The research training sites have included 80 universities, research laboratories, and Federal institutions in 34 states and five foreign countries. The home institutions are broadly representative, including 74 universities in 25 states and the Commonwealth of Puerto Rico.

There has been little change in the level of support for the Visiting Scientist Award Program. Only 10 awards have been provided, based on 17 applications reviewed. The requirement for considerable advance planning by the host institution and the prospective Visiting Scientist may have tempered a fuller usage of this award. However, concerted efforts to advertise and encourage applications for the Visiting Scientist Award are being made by program staff. The MARC Review Committee and the National Advisory General Medical Sciences Council strongly support continuation of this award mechanism.

To date, 223 applications for the Honors Undergraduate Research Training Grant Program have been received, of which 203 have been approved; 60 minority institutions have received awards, 11 of which were second 5-year awards and eight of which were third 5-year awards. In 1989, the honors undergraduate program had 165 graduates, bringing the total number of graduates since 1978 to over 1,450. To date approximately 80 percent of these graduates have gone to either graduate or professional schools.

EVALUATION

In a report issued in September 1988 by the congressionally established Task Force on Women, Minorities, and the Handicapped in Science and Technology, it was stated that "of Federal programs established to give minorities and women access to science and engineering, we found the Minority Access to Research Careers (MARC)... closest to what we need today. MARC is a prime example of a successful Federal intervention program."

While this assessment of MARC is most reassuring, evaluation of the improvements and accomplishments of the various components of the MARC Program remains a continuing effort of program staff. The scientific community, the MARC Review Committee, the National Advisory General Medical Sciences Council, and special Institute consultants cooperate in this effort through continued interaction with program and other NIGMS staff.

Significant accomplishments of the Program include the following:

The "honors" concept of the MARC Program has encouraged institutions that are recipients of awards to establish a different and more rigorous program for students in the basic sciences. Many have incorporated an honors thesis requirement for graduating MARC students.

The provision of funds for seminars and lectures, release time for faculty members, up-to-date equipment and supplies, as well as the addition of full-time and part-time faculty in disciplines not formerly represented have all contributed to academic enrichment. In addition, specialized courses have been set up to prepare students to take standardized tests-- an area where minority students have traditionally been weak.

The scientific activities of grantee departments have been strengthened, and as a result, it has become more common for MARC faculty members and trainees to be invited to present papers at local, regional, and national scientific meetings. We are beginning to see a larger number of students serving as coauthors of scientific papers.

The MARC Program, with its emphasis on research, has produced a new awareness of the possibilities available in biomedical research among students historically motivated towards careers in medicine, dentistry, and education. An increasing percentage of the trainees are entering graduate schools.

The fact that MARC students are performing well in graduate schools has provided an increased number of "role models" for minority students, and has resulted in a significant increase in the number of science majors at minority institutions having MARC awards.

A number of the graduates of the Honors Undergraduate Research Training Grant Program who were supported in the first years of that award have now received the doctorate degree. From information available, they have done extremely well and are viable candidates for postdoctoral support at prestigious institutions.

At the annual MARC Scholars Conferences, trainees are recruited by dozens of major universities and corporate entities. There is heated competition for MARC honors students among the large number of schools and private companies that run summer training programs. Such willingness to invest in the short-term training of these students is seen as a very positive indication of the quality of the MARC trainees.

A MARC Honors Undergraduate Scholar has recently been named American Indian Student of the Year by the National Indian Education Association.

Other Federal agencies, including the National Science Foundation and the Department of Energy, are now establishing programs, modeled on MARC, in the fields of physical science and engineering.

SUPPLEMENTAL ACTIVITIES

In November of 1988, the Program sponsored the Seventh Annual MARC Scholars Conference and Program Directors Meeting; similarly, an Eighth MARC Conference was held on October 31 through November 1, 1989. The 1989 conference, held at the Hyatt Regency Hotel in Bethesda, Maryland, was attended by more than 300 MARC-supported students. Most had prepared oral or poster presentations on their research and, in addition to discussing their scientific activities, spent a good part of the

3-day meeting interacting with graduate school faculty and scientist representatives of the National Institutes of Health to discuss opportunities in graduate education. They received information on the graduate school application process, the availability of financial aid, and expectations regarding the Graduate Record Examination. As part of the conference, participants attended scientific seminars and toured selected NIH research laboratories.

On the first morning, the students and their program directors were welcomed to the conference by Dr. Ruth L. Kirschstein, Director of the National Institute of General Medical Sciences. The keynote address was given by Dr. William Delauder, President of Delaware State College, and later that day a luncheon address, entitled "Mutations of Insulin - Effects on Processing and Secretion," was presented by Dr. Lydia Villa-Komaroff of the Harvard Medical School.

In keeping with the scientific theme of the conference, the first day's plenary session was a symposium, "The Structural Biology of Molecules." Symposium speakers were Dr. Paul Sigler of Yale University and Dr. Robert Sauer of the Massachusetts Institute of Technology.

On the second day, the MARC scholars were greeted by Dr. Louis Sullivan, Secretary of the Department of Health and Human Services. Following Dr. Sullivan's talk, a second symposium, "The Structural Biology of Viruses," featured Dr. James Hogle from the Research Institute of Scripps Clinic and Dr. Manuel Navia of Merck Sharp and Dohme Research Laboratories.

On the final day of the conference, a concurrent MARC Program Directors Meeting was held to provide a forum for the improvement of program planning and coordination of the MARC Honors Undergraduate Research Training Grants. Discussions centered on a number of issues such as requirements for renewal applications, requirements for predoctoral applications, training for faculty, summer research training programs, the impact of MARC on the institutional setting, the use of the visiting scientist program, pre-MARC student activity, and grants management issues affecting each of the 52 MARC Program Directors in attendance.

MINORITY ACCESS TO RESEARCH CAREERS (MARC) PROGRAM

Professional Staff

Elward Bynum	Director
Dolores Lowery	Program Administrator

Support Staff

Eileen Haskins	Secretary
Tobey Moxley	Grants Clerk

PHARMACOLOGICAL SCIENCES PROGRAM

SCOPE

This report reviews the activities of the Pharmacological Sciences (PS) Program for Fiscal Years 1988 and 1989.

PROGRAM OBJECTIVES

The ultimate goal of the PS Program is the support of research and research training leading to increased understanding of the interactions of drugs with living systems in order to produce new, safer, and more efficacious therapeutic agents. Research in the pharmacological sciences requires an integrated knowledge of the living organisms and the drugs themselves. Thus, the Program's area of coverage is highly multidisciplinary, supporting a broad range of research aimed at improving the understanding of biological phenomena and related chemical and molecular processes involved in the actions of therapeutic drugs and anesthetics. The scope of the Program's grant portfolio ranges from synthetic chemistry and molecular pharmacology to controlled studies in human subjects. Much of the research supported by the Program unites such fields as genetics, molecular biology, chemistry, and computer science with the more traditional areas of pharmacological investigation in an effort to characterize molecular interactions and to understand their expression in humans.

ORGANIZATION AND STAFFING

The PS Program currently consists of the Director, four program administrators, three grants clerks and the Secretary to the Program Director. Since the previous Biennial Report, Dr. M. Janet Newburgh has left the Program to assume the duties of the Institutional Liaison Officer in the Office of the Director, National Institutes of Health (NIH). Dr. Paul Velletri left in October, 1989, to join the Review Branch of the National Heart, Lung, and Blood Institute. Ms. Lisa Tetter left the Program to become a grants fiscal assistant in the National Institute of General Medical Sciences (NIGMS) Budget Office. Dr. Michael Rogers, formerly Executive Secretary of the Bioorganic and Natural Products Chemistry Study Section, Patrice Molnar, a former member of the US Navy, and Charlene Myron, formerly of the private sector, have joined the Program. Dr. Yvonne Maddox, Program Administrator in the NIGMS Biophysics and Physiological Sciences Program, and Dr. Warren Jones, Program Administrator in the NIGMS Cellular and Molecular Basis of Disease Program, have joined the PS Program on a part-time basis.

Dr. Carrico, in addition to directing the program, handles research grants and the postdoctoral training grants in clinical pharmacology; she also administers the Pharmacology Research Associate Program. Dr. Kuether administers research grants in biorelated chemistry and serves on the Fellowship Overview Group of NIGMS. Dr. Rogers also administers grants in biorelated chemistry and is responsible for the individual fellowships in the Program. Dr. Maddox and Dr. Jones handle portfolios of grants in pharmacology and biorelated chemistry, respectively. Note: The Program is currently recruiting for a replacement for

Dr. Velletri to handle grants in pharmacology and anesthesiology. The program administrators all administer institutional training grants.

FISCAL OVERVIEW

The PS Program supports research in three main areas: anesthesiology, pharmacology, and biorelated chemistry. It must be borne in mind that the basic criterion for support is excellence irrespective of the topic, and there is a great deal of overlap among these areas as befits the nature of the Program. Nevertheless, these categories are useful for administrative purposes. In FY 1988 and 1989, approximately 60 percent of the Program's research grant budget was expended for biorelated chemistry research, approximately 30 percent for pharmacology, and approximately 10 percent for anesthesiology. The figures in the fiscal summaries that follow reveal that funding for research training in terms of awards and number of trainees has remained relatively constant compared to the FY 1986-1987 period with some fluctuation in the number of individual fellowships. The Program has enjoyed modest growth both in terms of dollars and number of awards for research project grants, even during a period of fiscal restraint. We believe this reflects robust activity in the research areas supported by the PS Program.

PHARMACOLOGICAL SCIENCES PROGRAM

Program Area Summary
Fiscal Year 1988

RESEARCH		Awards	Dollars (In Thousands)
Research Grants			
Centers		5	$ 2,802
Program Projects		9	4,374
Research Projects		528	74,353
Total		542	81,529
Career Program			
Career Awards		1	33
Career Development Awards		1	54
Total		2	87
Conference Grants		1	2
TOTAL RESEARCH		545	81,618
RESEARCH TRAINING			
National Research Service Award			
Institution Fellowship (T32)	Trainee Positions		
o Predoctoral - Pharmacological Sciences	259	29	3,963
o Postdoctoral - Clinical Pharmacology	28	9	798
- Anesthesiology	10	3	319
Total Institutional Training	297	41	5,080
Individual Fellowships (F32)			
o Postdoctoral	67	67	1,414
TOTAL FELLOWSHIPS	364	108	6,494
Intramural NIH			
Pharmacological Research Associates	17		788
TOTAL RESEARCH TRAINING	381		7,282
R & D Contracts		1	12
TOTAL IN PROGRAM AREA	381	654	88,912

PHARMACOLOGICAL SCIENCES PROGRAM

Program Area Summary
Fiscal Year 1989

R E S E A R C H		Awards	Dollars (In Thousands)
Research Grants			
Centers		4	$ 2,779
Program Projects		10	4,828
Research Projects		536	78,221
Total		550	85,828
Career Program			
Career Awards		1	33
Career Development Awards		1	54
Total		2	87
Conference Grants		4	10
TOTAL RESEARCH		556	85,925
R E S E A R C H T R A I N I N G			
National Research Service Award Institution Fellowship (T32)	Trainee Positions		
o Predoctoral - Pharmacological Sciences	262	30	4,420
o Postdoctoral - Clinical Pharmacology	30	7	790
- Anesthesiology	10	3	349
Total Institutional Training	297	40	5,559
Individual Fellowships (F32)			
o Postdoctoral	46	46	1,025
TOTAL FELLOWSHIPS	343	86	6,584
Intramural NIH			
Pharmacological Research Associates	18		951
TOTAL RESEARCH TRAINING	361		7,535
R & D Contracts		1	10
T O T A L I N P R O G R A M A R E A	361	642	93,470

RESEARCH TRAINING OVERVIEW

Pharmacological scientists serve important roles in a wide range of scientific and health-related endeavors in academia, government, and industry. To provide a continuing source of such scientists is the purpose of the training programs of the PS Program. Since 1975, research training in the pharmacological sciences and clinical pharmacology has been supported by National Research Services Awards (NRSA's). Since 1983, research training in anesthesiology also has been supported with NRSA's under the auspices of the PS Program. Multidisciplinary predoctoral training programs in the pharmacological sciences are supported by institutional training grants, as is postdoctoral research training in anesthesiology and clinical pharmacology. Individual fellowships also are available for postdoctoral training in all of the above fields.

Predoctoral Awards-

The PS Program provides support for predoctoral training through the institutional grant mechanism. An important aspect of these awards is the multidisciplinary nature of the training. Institutional predoctoral fellowship programs in the pharmacological sciences provide students with a combination of skills and techniques necessary to address a wide spectrum of cutting-edge pharmacological problems. These programs are designed to offer students the opportunity to apply knowledge of molecular biology, medicinal chemistry, neurobiology, biochemistry, physiology, and other related sciences to problems of pharmacological interest. Regardless of their departmental affiliation, all trainees receive a thorough foundation in pharmacological principles and their application to biomedical research. In FY 1988, there were 29 institutional predoctoral training awards supporting 259 trainees. In fiscal 1989, there were 30 institutional training awards supporting 262 trainees.

Postdoctoral Awards-

Institutional postdoctoral fellowship programs in clinical pharmacology and anesthesiology are designed to train physicians in the application of laboratory research principles to the study of these disciplines. In FY 1988, there were 9 clinical pharmacology training awards supporting 28 trainees and 3 anesthesiology training awards supporting 10 trainees. In FY 1989, there were 7 clinical pharmacology training awards supporting 30 trainees and 3 anesthesiology training awards supporting 10 trainees.

Individual postdoctoral fellowships are supported in all areas, including bio-related chemistry, pharmacology, and anesthesiology. The PS Program supported 67 individual postdoctoral fellows in fiscal 1988 and 46 fellows in 1989. Approximately 80 percent of the individual fellowships awarded through the PS Program were in bio-related chemistry, reflecting the fact that most of the applications for individual fellowship awards which are received and are eligible for funding by the PS Program occurred in this scientific area.

Research Career Development Awards-

The PS Program accepts applications for Research Career Development Awards (RCDA) in all research areas relevant to its mission. The purpose of these awards is to assist exceptionally promising junior investigators in making the transition to established, independent investigators by providing salary support to free up additional time for research and training for the applicant. It is still relatively recently that the PS Program began to accept these awards from candidates with commitments to areas other than clinical pharmacology and anesthesiology. In FY 1988, one RCDA was supported by the PS Program. In FY 1989, one RCDA was supported.

Pharmacology Research Associate Program-

The Pharmacology Research Associate (PRAT) Program is the only intramural activity of NIGMS. This unique program provides opportunities for 22 highly qualified postdoctoral fellows to spend 2 years in intramural laboratories of the National Institutes of Health and those of the Alcohol, Drug Abuse, and Mental Health Administration. There are currently 116 formally appointed preceptors whose interests include pharmacokinetics, hormone-receptor interactions, cyclic nucleotide biochemistry, molecular toxicology, and virtually all aspects of neurosciences. In addition to the extramural support for clinical pharmacology, the PRAT Program provides an individualized curriculum for interested and qualified physicians to pursue research training in clinical pharmacology at the National Institutes of Health.

Minority Research Supplement Awards-

There has been an increasing awareness that to meet the scientific manpower needs of tomorrow we must begin today to more fully take advantage of the underrepresented minority pool of potential scientific talent. Thus, the National Institutes of Health has begun (see the April 22, 1989, issue of the NIH Guide to Grants and Contracts, Vol. 18, No. 14) a program that provides for administrative supplements to be made to existing NIH grants for the support of underrepresented minority undergraduates, graduate students, and minority investigators. In FY 1989, the PS Program received and made one award for the support of minority graduate students.

MERIT AWARDEES

The National Institutes of Health in July, 1986, initiated the Method to Extend Research in Time (MERIT) Award. It was intended to relieve the administrative burden of frequent grant renewal and enhance continued creativity through long-term support of outstanding investigators. These investigators are deemed to have demonstrated research competence and productivity in such a superior fashion that there is little doubt that they will continue their outstanding performance.

In NIGMS, nominations are considered once a year by the National Advisory General Medical Sciences Council; investigators do not apply for these awards. An investigator is considered for a MERIT award when the grant for his major research effort has been renewed the previous fiscal year. When the MERIT award is granted, the investigator has 5 years of support and can apply for renewal at the appropriate time with an abbreviated application that undergoes council, but not study section, review.

Currently, approximately 2 percent of NIGMS grantees hold MERIT awards, and for the 1988-1989 period, the Pharmacological Sciences Program had 10 MERIT awardees.

"New Methods for the Synthesis of Polynucleotides"
R37 GM 10265 (Letsinger, R.), Northwestern University

Dr. Letsinger's research is directed toward synthesis and characterization of modified oligonucleotides which retain the natural base as "recognition sites" for DNA or RNA but have alterations in the backbone that convey useful new properties. The general objective has been to build a chemistry base for rational design of oligomers that are superior in specific applications to the natural type oligonucleotides. The highlights of the current program are development of two new families of oligonucleotide analogs: the cationic oligonucleotides and the cholesteryl-conjugated oligonucleotides. The former have potential as unique probes in biological studies and diagnostic medicine. The latter open a promising direction for research on antiviral agents.

Cationic Oligonucleotides: These compounds differ from unmodified oligonucleotides in that they bear positive charges at the internucleoside links in place of negative charges. They are prepared by oxidative coupling of diamines with synthetic oligonucleotide hydrogen phosphonate derivatives. In aqueous solution at appropriate pH the terminal amines in the internucleoside aminoalkylphosphoramidates are protonated to afford a polycation. The salient features of this class of compounds are: (1) they form duplex structures with natural type anionic polydeoxyribonucleotides in aqueous solution, as shown by thermal dissociation data; (2) circular dichroism studies indicate that base stacking in these heteroduplexes is similar to that in the natural duplexes; (3) the stability of the cationic-anionic heteroduplexes increases with decreasing ionic strength, the reverse of the behavior of natural duplexes; (4) at low salt concentration (low ionic strength) the stability of the cationic anionic duplexes is greater than that of the natural duplexes; (5) stability of the cationic-anionic duplexes is strongly dependent on proper base pairing; and (6) dependence of the heteroduplex stability on salt concentration can be modulated by controlling

the ratio of negative and positive sites on the probe. These properties define a new self-organizing system in which the interaction is subject to selective tuning and control.

Cholesteryl-conjugated Oligonucleotides: Dr. Letsinger and coworkers have prepared and investigated the properties of oligonucleotides with a cholesteryl group tethered at the 3′-terminal internucleoside link (O=PNHCH$_2$CH$_2$NHCO-cholesteryl), the remaining backbone internucleoside links being conventional phosphodiesters (O=PO) on one hand and phosphorothioates (O=PS) on the other. This hydrophobic modification was introduced to favor interaction of the anionic oligomers with cell membranes, which are rich in cholesterol. They found that the cholesteryl-modified oligomers form complexes with complementary polynucleotides that are about as stable as those formed by the parent oligomers. Samples of these cholesteryl derivatives were tested as inhibitors of HIV (inhibition of syncytia formation and expression of viral proteins p17, p24, and reverse transcriptase in Molt-3 cells) by Dr. Prem Sarin at the National Cancer Institute. A striking result is that the cholesteryl modification significantly enchances activity of both the phosphodiester and the phosphorothioate oligomers. In the most favorable case, with a 20-mer cholesteryl-phosphorothioate derivative, complete inhibition in all assays was obtained at an oligomer concentration of 0.2 μM, less than 1/10 the amount of the corresponding unmodified phosphorothioate required for comparable inhibition. The relative enhancement was even greater with shorter oligomers. These results open new leads for designing antiviral agents and raise challenging questions concerning the mechanism of action of the modified oligonucleotides.

"Synthesis and Reactions of Peptides"
R37 GM 13453 (Kemp, D.), Massachusetts Institute of Technology

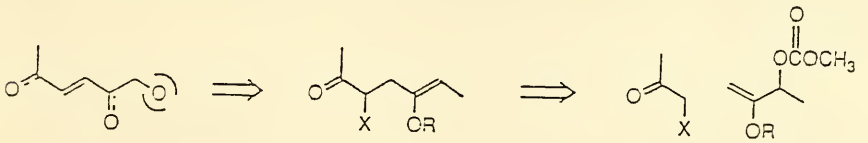
Significant progress has been made toward refinement of the thiol capture methodology and its application to the synthesis of medium-sized proteins of unnatural structure. The approach relies on the fragment condensation tactic, in which sections of backbone structure are prepared independently and then linked together in final synthetic steps. Using conventional linkage chemistry in which fragments are assembled in dilute solution and amide bonds are formed in the same step, reactions must be carried out in solvents other than water and with blocking groups attached to all reactive side chain functionalities. In the final step of a classical chemical synthesis of a protein, all blocking groups must be removed from as many as 50-60 side chains of the fragile protein molecule, and this step is usually prone to generate impurities that are extremely difficult to remove from the desired synthetic protein.

The thiol capture strategy differs from conventional linkage chemistry in that the amide bond formation step that permanently joins fragments follows a step in which a temporary sulfur-sulfur bond is formed. Amide formation thus occurs within a molecule, rather than between a pair of molecules, and the amide forming step is facilitated by proper design of a spacing element that positions the reactive partners in an orientation favorable to reaction. The initial sulfur capture step was chosen in part because it is optimally carried out in aqueous solutions that facilitate fragment solubility.

Three important advances with thiol capture have been realized during the past year. First, new blocking groups such as the N-Boc-2,2-dimethylthiazolidone-carboxamides have been developed to facilitate the sulfur capture step and spacing elements that achieve one hundred-fold selectivity between L- and D-cysteine residues have been demonstrated as a result of rational design. Second, a cheap, computer-controlled device has been constructed for automating the tedious and repetitive steps in solid phase peptide synthesis that are required to generate fragments compatible with thiol capture. Third and most important, two thiol capture couplings have been carried out, a coupling of 8 + 21 amino acid fragments from the structure of the small protein BPTI and a coupling of 13 + 12-amino acid fragments from the sequence of the protein ROP. In the former, only cysteine residues were blocked in the 21-mer, and in the latter, no amino acids bore side-chain blocking groups. In both cases, trouble-free synthesis was observed with yields of isolated, purified products in the range of 70 percent. These results imply the feasibility of semisynthesis by thiol capture, in which some fragments are derived directly by cleavage of native proteins. Hitherto, semisynthesis has only been achieved in special cases and with difficulty. An enormous saving in the labor of construction of unnatural proteins could be achieved if most of the amino acid sequence could be derived in this manner from cleavages of natural proteins.

"Synthesis of Macrolides, Steroids, Cyclopentanoids"
R37 GM 13598 (Trost, B.), Stanford University

The broad biological applications of macrocyclic natural products from antibiotic to antitumor to immunosuppressant activity make their de novo synthesis and selective modification high priorities. Dr. Trost and his coworkers have been engaged in developing novel catalysts that they call "the chemists' enzymes" to provide molecular sculpting and tailoring with high selectivity in order to create such rings with high efficiency. In developing relevant methodology, a recurring structural feature is depicted in structure 1. The enol ether 2 represents an ideal precursor since addition of a proton initiates elimination of HX to generate the desoxy series (lacking the oxygen in parentheses); whereas, addition of OH^+ initiates the same sequence to provide the full oxidation pattern. This pivotal intermediate derives from a

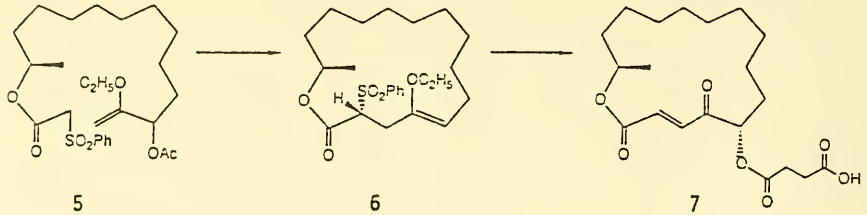


1 2 3 4

template directed formation of a carbon-carbon bond between precursors 3 and 4, which permit alkylation under totally neutral conditions.

To test this idea, the simply derived acyclic substrate 5 was exposed to a Pd organometallic complex, which provided the 16-membered ring macrolactone 6.

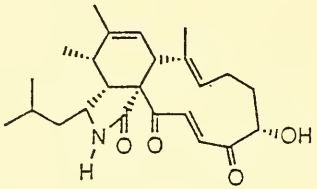
Introduction of the OH^+, followed by succinoylation, completes the synthesis of the antibiotic A26771B. The high stereocontrol for the introduction



5 6 7

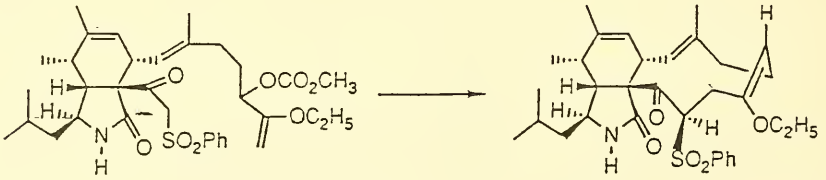
of the hydroxyl group arose because of the high conformational control exerted on the 16-membered ring of 6 by the phenysulfonyl moiety.

The cytochalasins constitute a more exciting challenge, in part because their varied biological activities have made them valuable tools in cell research and in part because of their higher structural complexity. Of this family, this group has focused on an [11]-cytochalasin; aspochalasin; B, 8; since the macrocyclic ring could not be constructed by a lactonization method.



8

Towards this goal, the acyclic substrate was synthesized from a readily available amino acid building block in a fascinating series of reactions that have extended our knowledge base for several important processes exemplified by the very important Diels-Alder reaction and generated new chemistry exemplified by a novel Claisen rearrangement. Most importantly, the critical template directed macrocyclization provides the 11-membered ring as a single entity.



9 10

The unraveling of the macrocycle to both desoxyaspochalasin B and aspochalasin B proceeds efficiently. Once again the phenylsulfonyl group plays the role of a stereochemical relay element, a new concept in synthetic strategy that may be of general importance. This route provides the most flexible and efficient entry to the cytochalasin family yet designed. Most significantly, the success of these syntheses validates the general strategy: This template directed reaction provides effective macrocyclizations by C-C bond formation and it also creates a versatile juxtaposition of functionality for final manipulation to targets.

"Biochemically Significant Molecules as Ligands"
R37 GM 13638 (Taube, H.), Stanford University

Dr. Taube's laboratory has recently begun to uncover rich organometallic chemistry of the moieties $Os(NH_3)5^{2+}$ and $Os(NH_3)5^{3+}$ which, by virtue of their composition, clearly belong in the realm of traditional coordination chemistry. It is a direct outgrowth of research into the basic chemistry of the congener element ruthenium which Dr. Taube began more than a decade ago and has depended for its success on having learned how to carry out reactions with traditional metal complexes in non-aqueous media.

The moiety $Os(NH_3)5^{2+}$ is unique in all of chemistry in forming robust bonds to benzene and other aromatic molecules. It attaches to adjacent carbon atoms of the benzene ring converting one of the three double bonds to a single bond, thereby greatly altering the reactivity of the remaining two and opening up novel opportunities in catalysis. For example, a transformation of potential usefulness is the selective hydrogenation of benzene to cyclohexene, which, at present, is accomplished at very low efficiency because the desired product, which has one remaining double bond, is more easily hydrogenated than benzene itself. When $Os(NH_3)5^{2+}$ attaches to benzene, the reactivity of the remaining two double bonds to hydrogenation is greatly enhanced, and conditions are easily found under which free benzene is not hydrogenated but the complex with the osmium moiety is. This moiety both activates for hydrogenation and protects the benzene from complete hydrogenation.

The affinity of the metal for the remaining double bond is so great that energy is expended in releasing the cyclohexene from the combination as described. As a result even though the process is highly selective, it does not constitute a practical solution to an important problem. It does constitute a novel approach and there is good reason to believe that by modifying the properties of the metal-containing unit, a practical process will result, appropriate not only for this particular process, but also for many other transformations of aromatic molecules.

The capacity to form robust combinations with molecules containing double bonds extends also to heterocyclic molecules such as pyridine, pyrrole, etc. (and derivatives), which are perhaps of more immediate biological interest than are the hydrocarbons. Dramatic changes in reactivity, which have been only partially explored, again ensue.

An advantage of the present system is that the nature of the metal-containing moiety can be changed readily and systematically. By the simple device of

working with a tetraammine rather than the pentaammine, the reactions of the metal center are dramatically altered. With benzene, a very stable complex results in which the metal is firmly bound to each of the six carbon atoms of benzene, the metal thereby losing one of four ammonia molecules. A number of amino acids contain the benzene ring, and there is every likelihood that a similar attachment will occur there. Whether this will occur for amino acids already incorporated into proteins remains to be seen. At the least, heavy metal marker atoms in robust combination can be introduced into synthetic polypeptides.

"Non-Toxic Thrombolytic Agent From Snake Venom"
R37 GM 15591 (Tu, A.), Colorado State University

During the course of comparing hemorrhagic and nonhemorrhagic proteases from snake venoms, Dr. Tu isolated atroxase, a nontoxic protease from Crotalus atrox (western diamondback rattlesnake) venom. Atroxase has a molecular weight of 23,500 and an isoelectric point of pH 9.6. Amino acid analysis indicates atroxase contains 206 residues with no sulfhydryl groups. Metal analysis found zinc and potassium at 1 mole/mole of protein. Atroxase is a fibrinogenase which cleaves the Aα chain of fibrinogen first and then the Bβ chain and shows no effect on the gamma chain. Atroxase also hydrolyzes fibrin. Fibrin solubilization occurs from the hydrolysis of the α-polymer and unpolymerized α and β chains. Although the crude venom induces platelet aggregation, atroxase demonstrated no ability to induce or inhibit aggregation.

In vivo, fibrinolytic activity of atroxase was tested on artificial thrombi induced in the posterior vena cava of Sprague-Dawley rats. Thrombolysis was then characterized by angiographic techniques over a period of 3 hours. Intravenous administration resulted in thrombolysis within 1 hour, followed by recanalization of the originally occluded vein within 2 hours. Fibrinogenolytic activity resulted in a 60 percent decrease in the rat's plasma fibrinogen level. Atroxase is nontoxic up to 20 μg/g in mice and is nonhemorrhagic in mice up to 50 μg/g. Histological examination indicates that atroxase is nontoxic to rat liver, heart, and kidney. This result indicates that snake venoms can be a good source for a nontoxic thrombolytic agent.

"Suicide Inactivation of Cytochrome P-450 Models: A Mechanistic Study"
R37 GM 17880 (Collman, J.), Stanford University

A characteristic feature of olefin epoxidation by cytochrome P-450 is its "suicide inactivation" via the formation of a kinetically incompetent β-hydroxy-N-alkyl hemin. In the process of catalyzing olefin epoxidation using oxygen atom transfer oxidants, model iron porphyrins exhibit the same suicide inhibition. For model iron porphyrins the regiochemistry and stereochemistry of the isolated N-alkyl porphyrin are the same as observed for the suicide adducts of P-450.

The efficiency of this inactivation was measured by a partition number, the ratio of the rate constants for epoxidation and hemin N-alkylation. The partition numbers of 1-alkenes (100-200) resemble those found for P-450.

Although the partition numbers are sensitive to the nature of the olefin and somewhat sensitive to the steric properties of the catalyst, they are relatively insensitive to the electronic properties of various hemins, and to the nature of several *p*-substituted styrenes.

There is little evidence for substantial charge separation during either process. These results are consistent with epoxide and N-alkyl hemin formation proceeding either by a concerted reaction or through a transient intermediate. The results of this study provide valuable insight into the mechanism of P-450 suicide inhibition by 1-alkenes.

"Biochemical Model Reactions"
R37 GM 18754 (Breslow, R.), Columbia University

The enzyme ribonuclease cleaves RNA in a two-step process: first it breaks the chain by forming a cyclic phosphate ester, and then water cleaves that cyclic phosphate. Dr. Breslow's laboratory has been studying chemical systems that mimic these two steps, and has recently discovered that they operate by a mechanism that was not previously suspected. This new information has permitted the design of much more effective catalysts that imitate the enzyme, and has probably furnished insight into the mechanism used by the enzyme itself.

In one of the studies, Dr. Breslow and his coworkers showed that the catalytic groups of the enzyme, imidazole rings, could catalyze the cleavage of RNA by themselves, although very slowly compared with the enzyme rate. The simple chemical system showed many of the characteristics of the enzyme system, particularly in the dependence of the rate on the pH of the medium. There were two possible mechanisms to explain the results of that early work; recently it has been shown which one is correct. The surprising finding is that the acidic catalytic group acts in the first step, putting a proton on the phosphate group of the RNA; it had been believed that in the enzyme it acts later to protonate one of the sugar oxygens. However, Dr. Breslow and his coworkers have reviewed the available information on the enzyme, and believe that there is much evidence that the enzyme also uses the mechanism discovered for this model system.

About 10 years ago an artificial enzyme had been constructed in the Breslow laboratories to perform the second step of the RNA cleavage sequence, hydrolysis of the cyclic phosphate. The catalyst had a binding cavity to hold the substrate, and two imidazole rings to perform the reaction. The design of this catalyst was based on the previous ideas about mechanism, but the new evidence suggested that the placement of the imidazole rings should be changed so the acidic catalyst group could more easily reach the phosphate oxygen. This has now been done, and the new catalyst is much more effective than were the previous versions. This confirms the mechanistic idea derived from other evidence, since the new catalyst geometry cannot use the original mechanism. Furthermore, the new catalyst is so effective that it cleaves its substrate only 230 times more slowly than the enzyme ribonuclease A cleaves its substrate. The results are a striking indication of the utility of mechanistic information in the design of enzyme mimics.

"Biological Studies With Liver Peroxisome Proliferators"
R37 GM 23750 (Reddy, J.), Northwestern University

Hypolipidemic drugs such as clofibrate, as well as certain plasticizers, pesticides, and herbicides, markedly increase the number of peroxisomes in liver parenchymal cells. These agents have been designated as peroxisome proliferators. Continued exposure to these peroxisome proliferators produces essentially a similar pleiotropic response leading eventually to the development of liver tumors in rats and mice. However, unlike a majority of chemical carcinogens, the carcinogenic peroxisome proliferators are persistently undetectable as mutagens with the short-term in vitro tests currently in use. Dr. Reddy and coworkers have postulated that events leading to or associated with the induction of peroxisome proliferation play a major role in the development of liver tumors by these non-mutagenic carcinogens. It was proposed that the carcinogenicity of these agents is mediated by oxidative DNA damage resulting from persistent peroxisome proliferation and over 20- to 30-fold increase in hydrogen peroxide-generating peroxisomal β-oxidation system enzymes in liver. Work from Dr. Reddy's laboratory demonstrated increased lipid peroxidation, excessive accumulation of lipofuscin and significant increases in the formation of 8-hydroxydeoxyguanosine in liver DNA of rats chronically exposed to peroxisome proliferators. These observations clearly establish that persistent proliferation of peroxisomes leads to specific oxidative DNA damage and this may be the mechanism by which these non-mutagenic carcinogens exert their carcinogenic effects.

Recent work from Dr. Reddy's laboratory has focused on the molecular mechanism by which structurally diverse peroxisome proliferators induce essentially similar biological effects. Using the cDNA's for the peroxisomal β-oxidation system enzymes, it has been shown that the mRNA levels of the β-oxidation genes increase 20- to 30-fold in the livers of rats treated with peroxisome proliferators. This increase results from enhanced transcription of the corresponding genes. Dr. Reddy and his colleagues proposed that peroxisome proliferators affect transcription of particular genes in a tissue-specific manner by binding to specific recognition molecules in responsive cell types. They have purified a peroxisome proliferator-binding protein from rat liver by affinity chromatography. Using ciprofibrate, clofibric acid, and nafenopin as affinity ligands and eluting agents, they have identified a major protein with an apparent molecular weight of 70,000 under denaturing conditions of polyacrylamide gel electrophoresis. The physicochemical and biological properties of this peroxisome proliferator-binding protein are under evaluation.

In an attempt to investigate the induction and regulation of other peroxisomal enzymes, Dr. Reddy and his coworkers have recently isolated and sequenced a full-length cDNA for rat liver peroxisomal urate oxidase. This clone has 1,378 base pairs with an open reading frame coding for 303 amino acid residues corresponding to a molecular mass of 34,931 daltons. They have shown that urate oxidase mRNA is detected only in liver and that the mRNA content increases only about three-fold in the livers of rats treated with peroxisome proliferators. It appears that peroxisome proliferators differentially regulate different peroxisomal enzymes.

"Design and Synthesis of Peptide Mimics"
R37 GM 30759 (Bartlett, P.), University of California, Berkeley

The goal of Dr. Bartlett's program is to devise strategies for the design of small, non-peptidic molecules that can act as three-dimensional mimics of larger, biologically active proteins. One goal of this multifaceted project involves preparing mimics of the critical tripeptide unit of the α-amylase inhibitor, tendamistat. The triad of amino acid residues, Trp-18, Arg-19, and Tyr-20, constitute the crucial region on this 73-amino acid inhibitor. The structure of this inhibitor in solution as well as in the crystalline state is known, and it is seen that the amino acid triad adopts a well-defined conformation. This team has designed some rigid, tricyclic molecules which can position the appropriate sidechains in this conformation, and is currently engaged in the synthesis of these molecules. They have also synthesized a cyclic hexapeptide derivative, cyclo-[Phe-Ala-Trp-Arg-Tyr-(D)-Pro], which they believe may adopt a similar conformation; however this material does not appear to inhibit α-amylase. They have recently developed a software program, CAVEAT, that enables the search of the Cambridge Structural Database for molecules that can serve as starting skeletons for the design of inhibitors of this sort. With this program, Dr. Bartlett and his group have identified a cyclic hexadepsipeptide, enniatin-B, whose backbone conformation is essentially identical to that in the crucial region in tendamistat and are now undertaking the synthesis of the appropriately functionalized analogs of this molecule as alternative mimic designs.

Mimics of the active loop of serine protease inhibitors are also being pursued. The large and structurally diverse class of proteinaceous inhibitors of serine proteases adopt very similar backbone conformations in their "active-loop" regions, leading to the suggestion that a small-molecule mimic of such an active loop could also inhibit this class of enzymes. Dr. Bartlett's group has devised and synthesized such a molecule, a 12-membered lactone-lactam incorporating residues corresponding to the P2-P1-P1′ residues of human urinary subtilisin inhibitor (HUSI), which, despite its name, is an inhibitor of chymotrypsin. The synthetic mimic is also a chymotrypsin inhibitor, with a Ki value of 10μM. However, studies with control compounds suggest that the inhibitor is not optimized, and that its minimum energy conformation may not be the desired one. Second-generation designs are currently being pursued. With the aid of CAVEAT, a novel set of bicyclic pyridine derivatives as "third-generation" mimics of HUSI, and their synthesis is currently underway.

Finally, a recently initiated, long-range project is on interfacial competitors: inhibitors of ribonucleotide reductase. It is planned to devise mimics of oligopeptides that disrupt subunit association of multimeric enzymes. They have started with E. coli ribonucleotide reductase, an α,β-dimeric enzyme, and in initial work have synthesized octapeptides corresponding to the C-terminal region of the β-subunit. These peptides inhibit the enzyme, with Ki values in the mM range, and will be the subject of an NMR study (using the transfer NOE technique) to determine the conformation that they adopt when they bind to the complementary subunit. If it can be demonstrated that they are indeed functioning as anticipated and if their bound conformation can be determined, then conformationally constrained mimics

(perhaps using the α-helix or β-turn templates) will be devised as potentially more potent inhibitors of this enzyme. Other ongoing work involves the design of complementary bicyclic structures to serve as helix nucleators at the N- and C-termini of α-helices, bicyclic structures to induce β-turn conformation in a tetrapeptide loop, and conformationally constrained inhibitors of thermolysin.

"Heme Protein Structure and Function"
R37 GM 40168 (Peisach, J.), Yeshiva University

The major effort of Dr. Peisach's research has been directed toward explaining the properties and mechanism of action of the anti-tumor drug, bleomycin, especially the iron-containing form of the drug. The approach to this research has generally involved magnetic resonance spectroscopic techniques such as electron paramagnetic resonance (EPR) and electron spin echo envelope modulation (ESEEM) as well as analytical chemical and kinetic methods.

Three bleomycin studies were recently completed. A comparison was made of the products released from nuclei and purified DNA in reactions with activated iron bleomycin. Activated bleomycin is the name given to the form of the drug generated upon mixing Fe(II) with the commercial product, Blenoxane, in air. The activated form is a hypervalent iron-oxo species, formally equivalent to Fe(V). This first study showed that chromatin in intact nuclei was susceptible to damage by activated bleomycin and also that the nuclear membrane was oxidatively damaged in the reaction. The DNA-derived products from nuclei were structurally identical to those released from purified DNA though their yield and stoichiometries were different. It was also shown that ferric bleomycin has a high affinity for nuclei and gains access to nuclear DNA, a result implicating a role for iron in the in vivo function of the drug.

A second study demonstrated that t-RNA is susceptible to cleavage by activated bleomycin with a specificity related to the structure of this nucleic acid. Until this report, ribonucleic acid was thought to be resistant to cleavage by bleomycin. The result has important implications in the molecular pharmacology of this type of drug since the attack on a target other than nuclear DNA could be partly responsible for its toxicity.

A third study was directed at the detailed understanding of the oxidation state of iron in bleomycin upon iodosobenzene activation of the ferric drug complex. The hypervalent iodine reagent has been used as an oxygen donor in heme protein (peroxidase) and bleomycin activation reactions with organic substrates. The results of this study showed a rapid modification of the drug molecule accompanied by removal of Fe(III) from the Fe(III)-bleomycin complex in the presence of iodosobenzene. No spectroscopic evidence could be found for activated bleomycin or a species like the intermediate, compound I, of peroxidase, in mixtures containing Fe(III) bleomycin and iodosobenzene. Fluorescence spectroscopy demonstrated a rapid modification of the bithiazole moiety of the drug upon reaction of either metal-free or ferric bleomycin, suggesting that the iodosobenzene reagent acts directly on the organic substituents of the drug without implication of an activated iron bleomycin intermediate.

PHARMACOLOGY RESEARCH OVERVIEW

The PS Program's grant portfolio in pharmacology represents a wide array of newer and more traditional areas of research in pharmacology, a blend that represents both the fundamental underpinnings and the evolution of the field. Research ranges from very basic areas such as molecular pharmacology of receptors to clinical studies in humans. Specific areas include the physiological disposition of drugs, the molecular basis of drug action, drug metabolism, factors that modify drug action, and toxic effects of therapeutic agents. While other NIH institutes support organ- or disease-targeted pharmacology, the research supported by the PS Program of NIGMS tends to be basic and general studies of pharmacology, of universal interest to the pharmacological community.

As with many other areas of biomedical research, the field of molecular pharmacology is making use of biophysical and molecular biological techniques to elucidate structural characteristics of macromolecules of pharmacological interest. Many investigators supported by the PS Program have been defining the structure of proteins that act as receptors or enzymes, or that function in transmembrane signaling, in the hopes of understanding not only how drugs exert their activity but how these macromolecules behave under normal homeostatic conditions in the cell.

Remarkable structural similarities are being discovered among many of the proteins that are embedded within the cell surface membrane. One outcome of these observations is that researchers are considering a new pharmacological taxonomy for receptors based on physical instead of solely functional properties of receptors. In fact, it is quite possible that a new appreciation for interrelationships among receptors and other integral membrane proteins, such as ion channels, may result from these new research directions.

Recently, considerable attention has been focused on two basic structural patterns in membrane proteins, in large measure because these patterns seem to appear so frequently. One pattern, found in receptors that are coupled to transducing elements within the membrane known as G proteins, consists of seven membrane-spanning regions. A second pattern, ordinarily associated with ion channels, consists of six membrane-spanning regions. Not all membrane proteins, of course, fit into these two structural motifs, but there are a large number of proteins of pharmacological interest that are encompassed by these two patterns.

One example of the new emphasis on structural approaches to pharmacological problems is the work of investigators who have been studying the structural and molecular properties of various components of the acetylcholine system. Studies have been performed to help clarify the roles of specific subunits of the nicotinic acetylcholine receptor in eliciting a cellular response. The nicotinic receptor has a structural motif that is characterized by five subunits with a central ion channel. By using circular dichroism and proton nuclear magnetic resonance to measure the binding of various toxins to short synthetic peptides that correspond to portions of the alpha subunit of the

nicotinic receptor, it has been possible to assign functional properties to specific subunits. Other studies have been performed that clearly identify the binding site of local anesthetics to the gated sodium channel that is present at the center of the nicotinic receptor.

Other investigators have interests in the muscarinic acetylcholine receptors that also transmit the chemical signal of the acetylcholine molecule to the interior of the cell. An important area that requires continuing clarification is the classification of muscarinic receptor subtypes not only by the classical pharmacological profiling, but also through molecular biological techniques. Through a series of cloning studies at a number of laboratories throughout the United States, scientists have identified at least four subtypes of muscarinic receptors and have come to the fascinating conclusion that these receptors have a structural affinity not with their functionally close cousin the nicotinic receptor, but with the other receptors that possess seven membrane-spanning regions and that bind to the large family of so-called G proteins. By developing antibodies against unique sequences of oligopeptides obtained from cDNA clones, investigators have been able to detect and map a number of muscarinic receptor subtypes.

Molecular biological techniques have also led to studies that have demonstrated a clear structural connection between the enzyme, adenylyl cyclase, which catalyzes the formation of an important second messenger, and sodium and calcium channels, suggesting a common genetic origin for these functionally diverse proteins. Ion channel proteins share a different structural motif than observed with any of the receptors described above. To date, all the ion channels studied in detail, i.e., the sodium, potassium, and calcium channels, possess a basic motif of six transmembrane regions. The structural studies recently performed on adenylyl cyclase open interesting new possibilities about the evolution and possible functions of this enzyme.

Important advances are also being made in traditional areas of pharmacology. The development of novel methods for drug delivery and studies of factors controlling drug absorption/distribution continue to be areas of considerable research activity. Studies on the controlled delivery of drugs and targeting of drugs to specific organs are advancing. Factors in drug absorption and the role of plasma binding are being studied, as are improved methods for data analysis for pharmacokinetics. Renal handling and transformation of drugs remain active areas.

A significant amount of research in the pharmaceutical sciences has focused on modifiers of drug action. These investigations have sought to predict the biological consequences of exposure to xenobiotics and to clarify individual differences in susceptibility to these substances. As a result of such studies, new information has come forth regarding the molecular mechanisms responsible for the induction of specific enzyme systems after exposure to toxic agents. The enzyme system most widely studied is the cytochrome P-450 enzyme family, which is responsible for the biotransformation, in the liver, of endogenous and toxic substances. Pharmacologists have made excellent progress in identifying the DNA sequences of mammalian cytochrome P-450 genes and their tissue specificity. Studies are also under way to determine the structural changes which accompany the developmental, as well as the drug-induced activation, of the P-450 genes. Other studies on P-450 enzymes have focused on delineating mechanisms for the individual enzymes. Metabolic

studies have indicated possibilities for altering metabolism by altering the chemical structure of the drug. Furthermore, studies of individual toxins and mechanisms of toxicity are revealing potential new therapeutics and new fundamental knowledge about the cell under stress.

Clinical pharmacology is an area which crosses all the institutes of the NIH. While the primary emphasis in clinical pharmacology in NIGMS continues to be research training, there is also a substantial amount of support on problems in clinical pharmacology which are generally applicable to a wide range of drugs. In many cases this support provides the research milieu in which research training programs are centered. Pharmacokinetics continues to be a vigorous area, especially in the area of modeling. There has been an increase in support for research targeted towards understanding the basis for individual responsiveness to drugs. Another critical area that receives strong support from the PS Program is the study of drug-drug and drug-disease interactions with the ultimate goals of both predicting and preventing such interactions. These areas are only a few of the many important areas in which the PS Program supports research in clinical pharmacology.

Inquiries such as those outlined above open up new possibilities for the discipline of pharmacology. By utilizing approaches from many different fields, pharmacological research is poised to answer recurrent questions that were perhaps unanswerable as little as a decade ago.

"Drug Macromolecule Complex Formation"
R01 GM 18360 (Taylor, P.), University of California, San Diego

Dr. Taylor and his colleagues have been examining the molecular structure of acetylcholinesterase (AchE) and the organization of the gene encoding this enzyme. Since the mid-19th century, acetylcholinesterase inhibitors have enjoyed therapeutic applications in the treatment of glaucoma, smooth muscle atony, certain arrhythmias, myasthenia gravis, and to reverse competitive neuromuscular blockade. Recent reports of the use of AchE inhibitors in the amelioration of Alzheimer's disease have appeared and AchE is the target site of the widely used agricultural and garden insecticides.

The cloning of the AchE gene and analysis of the disulfide linkages in the enzyme have identified a new family of serine hydrolases which are distinct from the pancreatic or subtilisin classes of hydrolases. Of interest is that AchE is homologous to the C-terminal region of thyroglobulin and all of the intrasubunit linked cysteines are conserved between thyroglobulin and AchE. Thus, nature has employed the same molecular architecture to hydrolyze a neurotransmitter as to synthesize the critical metabolic hormones, thyroxine and triiodothyronine.

Acetylcholinesterase's gene structure reveals that a single gene encodes the enzyme and alternative mRNA processing gives rise to the molecular diversity in the cholinesterase species. The first 480 amino acids which include the essential catalytic residues are encoded in a single exon. The second exon of 45 amino acids completes the third intrasubunit disulfide loop. The third exon completes the open reading frame, contains an extended 3′ untranslated region, and is alternatively spliced. One of the alternative third exons gives rise to the asymmetric species which consists of multiple catalytic

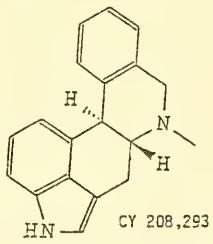
units disulfide linked to a collagen-containing tail unit. The other alternative exon encodes a sequence which is processed post-translationally by cleavage of a hydrophobic peptide and addition of a glycophospholipid on the C-terminal carboxyl residue. Hence, two separate mechanisms are available tethering this enzyme at distinct extracellular locations within the synapse (the basal lamina and the extracellular face of the plasma membrane). Unique inverted repeat sequences have been identified in the 3′ untranslated regions of the alternative exons which may play an important role in the exon selection.

Genomic cloning has also allowed the characterization ofthe 5′ untranslated region and flanking regions of the gene. Sequencing has uncovered potential regulatory sequences. Such promoter and enhancer regions may be important to the understanding the regulation of expression of proteins in the cholinergic nervous system.

"Clinical Pharmacology Center"
P01 GM 22220 (Murphy, M.), University of Chicago

The overall focus of this center is studying the function of the dopaminergic transmitter system in health and disease with particular emphasis on receptor function and the development of novel dopamine analogs.

During the characterization of peripheral dopamine receptors, this group found a very unusual structure represented by CY 208,293 showing DA_1 agonist activity. Previous studies have indicated that two OH groups analogous to the 3,4-OH groups of dopamine are essential for this agonist activity. It was, therefore, surprising to note DA_1 agonist activity in this compound which contained an ergoline nucleus with no phenolic or catecholic OH groups. Even more striking was the fact that this compound lacked apparent DA_2 agonist activity which has been represented consistently in the ergoline molecule. Understanding the significance of this finding will require additional studies. At this point, this finding represents a steep departure from the established structural requirements for DA_1 and DA_2 agonist activities.



CY 208,293

During the last 2 years this group has directed significant effort toward delineating the catecholamine (α-adrenergic and dopamine) receptors involved in the inhibition of ganglionic neuronal transmission. Experiments were conducted on the cardiac sympathetic ganglia of the dog. In the first study presence of DA_2 receptors in this ganglion was confirmed, but their results showed absence of DA_1 receptors, as had been proposed by other investigators.

In a subsequent study the presence of α_2-adrenoreceptors in addition to DA_2 dopamine receptors was noted. Furthermore, it appeared that α_2-adrenoreceptors may be physiologically more important than any other receptor. If dopamine is the transmitter of this modulatory effect, as has been suggested in previous studies, this work suggests that the receptor mediating this effect is the α_2-adrenoreceptor.

During the period of this report a new compound, dopexamine, which has been proposed to be yet another approach to the treatment of congestive heart failure, was studied in this laboratory. These studies led to the discovery of an additional property in this compound; namely, the inhibition of norepinephrine uptake into the nerve terminals. This is a significant finding, as it could explain at least a part of the mechanism of action of this compound on the heart. However, this action of dopexamine also represented a potential problem in the clinical setting. These results have led to the distribution of a warning against the use of dopexamine in a particular clinical situation, i.e., contraindicated within the first 48 hours of myocardial infarction.

This group has also been conducting behavioral research with rats designed to establish the functional significance of the DA_1/DA_2 receptor distinction *in vivo* both in the normal and the diseased brain. Initially they developed a behavioral preparation that was both a sensitive and selective measure of DA_1 or DA_2 receptor activation. They used a drug discrimination paradigm in which rats were trained to make one behavioral response following an injection of a dopamine agonist and another response following an injection of saline. The assay showed the appropriate potency relationships when other agonists were tested and the effects of agonists were blocked only by antagonists of the same type. That is, the assay proved to be a sensitive and selective behavioral assay for studying the functional effects of central DA_1 and DA_2 dopamine receptors *in vivo*. Moreover, there results suggest that the DA_1 receptor in the central nervous system is similar to the DA_1 receptor in the periphery.

More recently, they have begun utilizing the assay to investigate the function of dopamine receptors *in vivo*. Rats trained to respond to quinpirole (QUIN), a DA_2 agonist, were given various doses of SKF 38393 (SKF), a DA_1 agonist, in combination with QUIN. SKF in combination with QUIN caused a dose-related shift to the left in the QUIN dose-response function, consistent with the suggestion that stimulation of DA_1 receptors can modify DA_2-mediated behaviors. The results obtained suggest that one function of DA_1 receptors in the brain is to modulate dopaminergically mediated behaviors.

Methylnaltrexone is a narcotic antagonist designed to remain peripheral to the blood brain barrier. Thus, it is an interesting compound to explore the peripheral effects of opioids and may have clinical utility in blocking undesirable narcotic side effects while not interfering with centrally mediated analgesia. The last year has seen the completion of two studies in this area. This laboratory has demonstrated the ability of methylnaltrexone to block morphine-induced emesis in the canine. This confirms the functional presence of morphine emetic triggers on the peripheral side of the blood-brain barrier. The peripheral aspects of narcotic cough suppression also have been examined. While there is a central cough control center, there has been

controversy as to the site of narcotic activity in suppressing cough. There studies have shown reversal of morphine-induced cough suppression in the guinea pig, again suggesting that this is an action peripheral to the blood-brain barrier.

"Corticosteroid Pharmacokinetics/Dynamics"
RO1 GM 24211 (Jusko, W.), State University of New York at Buffalo

The time course of diverse pharmacological effects produced by corticosteroids such as prednisolone usually exhibits a slow onset, a delayed maximum response, followed by a slow dissipation of the effect all over an 18-30 hour period. Efforts in Dr. Jusko's laboratory have sought to establish methods for quantitating the receptor-gene mediated mechanisms of steroid effects and to evolve comprehensive mathematical models for quantitating the kinetics and dynamics of these effects in animal and human systems.

An improved model describing receptor-gene-mediated pharmacodynamics of prednisolone was developed which consists of seven differential equations. Data for plasma prednisolone concentrations, free hepatic glucocorticoid receptors, and hepatic tyrosine aminotransferase activity (TAT) following various doses of prednisolone were used to quantitate the kinetics and dynamics of this synthetic steroid in the rat. Ex vivo measurements of steroid association and dissociation rates with receptors provide k_{on} and k_{off} values for use in the model. The model provides for a coupling and simultaneous fitting of receptor and TAT data and is able to describe the recycling of receptors between cytosol and nucleus and the return of cytosolic receptors to baseline at later times following glucocorticoid elimination from the body. A numerical technique to determine the efficiency of TAT induction based on area under the curve calculations (AUC) supports the hypothesis that nonlinear dose-response effects caused by corticosteroids are due to dose- and time-dependent receptor depletion in the cytosol. The AUC of steroid effects is generally proportional to the AUC of drug-receptor binding (receptor occupancy).

Simulations were performed to examine the major determinants of corticosteroid effects and to compare the effects of single- and multiple-dosing regimens in maximizing drug effects. Biological factors (receptor binding, mRNA, and protein synthesis and degradation) as well as steroid persistence in the body control the classical slow onset and dissipation of steroid effects. This system and model offers opportunities to assess dosage regimen, drug interaction, and disease state effects on corticosteroid kinetics, receptor binding, and pharmacodynamics.

"Selective Destruction of Cytochrome P-450 by Drugs"
RO1 GM 25515 (Ortiz de Montellano, P.), University of California, San Francisco

Dr. Ortiz de Montellano and his colleagues have continued their investigation of the mechanism and function of cytochrome P450 enzymes using substrate probes. In one such study, they were able to estimate the rate at which the

radical thought to be an intermediate in the cytochrome P450 catalyzed oxidation of hydrocarbons is trapped by the iron-bound activated oxygen ($k = 2 \times 10^{10}$). This rate is rapid enough to rationalize the fact that substrate radicals rarely escape from the active site of the enzyme. The rate was determined with the help of a "radical clock" substrate that yields a radical that rearranges at a known rate, so that the rate of the oxygen trapping reaction can be estimated from the ratio of the rearranged and unrearranged products. The very high rate of the oxygen trapping reaction required the development and use of a faster radical clock than previously available. This radical clock substrate may be useful in studies of radical intermediates in the reactions catalyzed by other hydrocarbon oxidases.

The development of isozyme-specific mechanism-based inactivating agents for cytochrome P450 isozymes has continued with a focus on the use of acetylenic substrates to inactivate fatty acid hydroxylases. Previous studies have demonstrated that the inactivation of some cytochrome P450 isozymes by acetylenes is due to alkylation of the heme group. A recent study of the inactivation of one isozyme by substituted phenylacetylenes has shown that inactivation, in contrast to metabolite formation, is insensitive to the electronic properties of the substituents. This indicates that inactivation and metabolite formation involve distinct reactions of the enzyme with the triple bond. Recent work on inactivation of hepatic lauric acid ω-hydroxylase by terminal acetylenic fatty acids has shown, however, that this enzyme is inactivated by a mechanism other than heme alkylation. It is likely that the enzyme is inactivated by reaction of a reactive metabolite, probably the ketene, with a protein residue. The absence of heme alkylation argues that the enzyme is constructed so that the oxygen can only be delivered to the terminal carbon of the triple bond since heme alkylation requires delivery of the oxygen to the internal carbon. Terminal acetylenic fatty acids of different lengths have been shown to differentially inactivate the plant isozymes that hydroxylate fatty acids at the terminal versus internal carbons. Addition of a hydroxyl group four carbons from the terminal acetylenic group in a long chain fatty acid structure has resulted, furthermore, in an agent that is highly specific for the inactivation of prostaglandin co-hydroxylases. Efforts are continuing to develop agents that retain their inactivating activity in vivo and to use these agents to explore the biological functions of the various classes of fatty acid (ω-hydroxylases).

A study of a group of compounds known as the sydnones, some of which have useful pharmacological activities, has led to the discovery of a new class of mechanism-based inactivators of cytochrome P450 enzymes. These agents are oxidized to diazonium or diazo metabolites that alkylate the prosthetic heme group of the enzyme. In some instances the resulting heme adducts are potent inhibitors of the terminal step of heme biosynthesis and cause the accumulation of heme precursors. This clarifies the earlier observation by others that at least one sydnone causes the accumulation of red fluorescing pigments in the livers of treated animals.

"Biochemical and Genetic Analysis of Yeast Ras Function"
RO1 GM 28310 (Bourne, H.), University of California, San Francisco

Dr. Bourne's laboratory has recently uncovered GTPase-inhibiting mutations of the α_s gene of human pituitary tumors. In order to convert an α gene into an

oncogene, the mutation theoretically should occur in a cell that is normally induced to proliferate by hormones that stimulate adenylyl cyclase. Growth hormone (GH)-secreting pituitary cells meet this requirement, in that they are stimulated to proliferate by GH releasing hormone, which acts via intracellular cAMP. A subset of GH-producing pituitary tumors exhibit constitutively elevated GH secretion, cellular cAMP, adenylyl cyclase, and G_s activity.

In each of four tumors with this biochemical phenotype, Dr. Bourne's laboratory has found an α mutation that inhibits GTP hydrolysis. The α_s mutation in one tumor is located in the glutamine-227 codon, just as predicted from the Gln→Leu phenotype already characterized. Each of the other three tumors contains an α_s gene with a mutation at a different position, arginine-201. Two mutations replace the arginine with cysteine, while one substitutes histidine. Arg-201 is the α_s residue that is ADP-ribosylated by the exotoxin of V. cholerae. Effects of the toxin-catalyzed covalent modification are virtually identical to those produced by mutational replacements of the same residue: Enhanced GTP-dependent stimulation of adenylyl cyclase and reduced k_{cat}.GTP. This functional similarity makes it quite likely that the side chain of arginine-201 interacts with or forms a part of the molecular timing device that regulates k_{cat}.GTP. All known G protein α chains contain an arginine residue at the position corresponding to arginine-201, suggesting that this residue may play a similar role in regulating GTPase in other α chains. These findings also raise the possibility that other α chain genes may contain oncogenic mutations.

"Heme Protein Reductases"
RO1 GM 31756 (Sligar, S.), University of Illinois

Metalloproteins play an important role in biochemistry, encompassing oxygen transport and storage by the heme protein myoglobin and hemoglobin, central metabolic functions of xenobiotic catabolism and steroid hormone biosynthesis by P-450 oxygenases, and oxidative turnover through electron transfer by cytochromes and copper proteins. Theory has generally dictated that study of the mammalian cytochromes with site-directed mutagenesis would be extremely difficult due to problems in expressing eukaryotic metalloproteins as holoenzymes in high yield with a bacterial host. By assembling genes for these protein systems from individual nucleotides, it was shown that a variety of metalloproteins could be expressed in E. coli to extremely high level as soluble holoproteins.

Dr. Sligar's laboratory has used the modern techniques of site-directed mutagenesis and gene expression, together with new applications of biophysical principles (such as high pressure) to examine the interrelationships between the structure and function of several important metalloproteins. They have utilized mutagenesis to completely change the catalytic activity of a protein or enzyme. In one case, a b-type electron transfer cytochrome was converted into a catalytic oxidative demethylase. In another case, myoglobin was converted to catalase by replacing axial metal ligands. Myoglobin has become cytochrome P-450 by adding a cysteine in place of histidine. This group also completely redesigned the active site of cytochrome P-450 via genetic engineering to drastically alter the regio- and stereospecificity of catalysis, thus realizing a long-sought goal of de novo enzyme design and the

understanding of xenobiotic metabolite specificity. In addition, they have cloned and mutated the classical four-helical bundle protein b_{562} with a goal of understanding protein folding and used totally synthetic and/or cloned genes of b_5, plastocyanin, sperm whale myoglobin, horse heart cytochrome c, and P-450 to alter those residues involved in electron transfer catalysis, thus providing the means for documenting path-dependent redox transport in these systems.

"Acetylcholine Receptor: Structure, Function, Biogenesis"
R01 GM 32629 (Hawrot, E.), Yale University

The nicotinic acetylcholine receptor (nAchR) found at the neuromuscular junction is perhaps the best characterized of all neurotransmitter receptors. As a ligand-gated ion channel protein it is part of a large super-gene family which presently includes the neuronal nicotinic receptors, the glycine and GABA receptors, and will probably also include the NMDA receptors of the central nervous system. Although the mechanism by which the binding of a neurotransmitter to its binding site is transduced into gating of the intrinsic ion channel is currently unknown, it is likely that a molecular understanding of these events will emerge from the continued detailed study of the nAchR. As a first step towards this long-term objective, Dr. Hawrot has been directing his efforts to the identification of the ligand binding site on the α-subunit of the nAchR. These studies have been facilitated by the use of monoclonal antibodies and synthetic peptide fragments of the α-subunit.

Since the binding site for curaremimetic neurotoxins such as α-bungarotoxin (BGTX) appears to overlap the acetylcholine (Ach) binding site, Dr. Hawrot has used this toxin as a probe of the ligand binding site because of its higher affinity for the nAchR. He previously showed that small peptide fragments extending from amino acid residue 173 to about 204 on the α-subunit retain the ability to bind BGTX, albeit with reduced affinity compared to the native receptor. With the assumption that the three-dimensional structure of the complex formed between the synthetic peptides and BGTX will be similar to that resulting from binding of BGTX to the native receptor, Dr. Hawrot has carried out a number of spectroscopic studies to determine the structure of the peptide-BGTX complex. These studies have included measurements of changes in intrinsic fluorescence and circular dichroism upon binding.

Due to the high relative content of fluorescent aromatic residues within the binding domain peptides, a marked enhancement in intrinsic fluorescence levels can be detected upon binding of these peptides to BGTX. This enhancement has been used as the basis for a solution binding assay. With this assay, Dr. Hawrot has determined that the affinity of BGTX for an 18mer peptide (residues corresponding to α181-196) is 0.3 μM, while the interaction with a 12mer (α185-196) has an apparent K_D of 1 μM. Furthermore, Dr. Hawrot has obtained evidence that at least a portion of the enhanced fluorescence is due to the presence of a tyrosinate anion in the synthetic peptide. This observation may have important consequences in determining the identity of the negative subsite within the binding domain responsible for interacting with the positively charged nitrogen found in all nicotinic drugs. Analysis of changes in circular dichroism suggest that there is a large increase in β-structure upon binding and that an aromatic residue, most probably TRP-28 in BGTX, becomes relatively immobilized upon binding. This finding is consistent

with the presumed importance of this residue in toxicity based on its invariance among a large number of related neurotoxins.

NMR studies of changes in the aromatic region of the proton spectrum have indicated major perturbations in a number of key aromatic residues in both the synthetic peptide fragments and in BGTX. In order to interpret these changes the complete sequential assignment of all of the main and side chain proton resonances in the 12mer was determined. Similarly, Dr. Hawrot assigned most of the resonances of the 18mer. The assignments for BGTX have been previously published. With this information, Dr. Hawrot was able to conclude that the C4 region of Trp-28 in BGTX is greatly affected upon binding as are both histidine rings (His-4, His-68). In addition, the histidine on the peptide fragment is also perturbed upon binding. The NMR studies suggest that these perturbed residues may play an important role in binding. Since there is presently little information about the spatial configuration of any receptor's ligand binding site, these studies provide an important advance in our ability to identify and characterize the role of functionally important residues within the nAchR's binding domain.

Since many drugs produce their effects by acting at natural receptor binding sites, the detailed structural knowledge of a receptor's ligand binding domain such as that being pursued in this program will provide a rational basis for future drug design. Information on the structure of a binding site can be used to direct the preparation of suitable chemical analogs specifically designed to satisfy the steric and chemical requirements of the binding site.

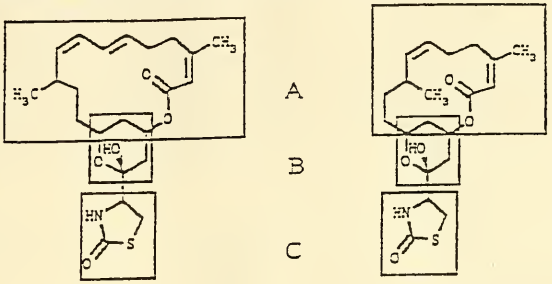
"Latrunculins - Chemical and Biological Aspects"
R01 GM 32776 (Spector, I.), State University of New York at Stony Brook

Actin is the most abundant protein in most eukaryotic cells. Much is known about the structure and biochemistry of actin and it is widely accepted that actin microfilaments play a major role in a whole range of cell movement phenomena. The highly dynamic nature of actin structures in nonmuscle cells makes it difficult to study in these cells the organization of actin microfilaments and to relate this organization to such key cellular processes as motility, shape determination, cell growth and division, and cell transformation. The use of specific drugs that interfere with the state of organization of actin *in vivo* and *in vitro* and affect cell behavior may overcome this difficulty.

Several years ago Dr. Spector and colleagues discovered that two new toxins from the Gulf of Eilat sponge *Latrunculia magnifica*, the latrunculins, induce profound changes in morphology and microfilament organization of nonmuscle cells. These initial results prompted Dr. Spector and colleagues to undertake a comprehensive investigation of the latrunculins that involved three interconnected efforts: 1) A chemical effort to determine the relationship between chemical configuration and biological activities. 2) A biochemical effort to elucidate the site and molecular mode of action of the latrunculins. 3) A cell biological effort to determine the biological and physiological consequences of the latrunculin action.

Structure elucidation of the first two toxins termed latrunculin A (LAT-A, 1) and latrunculin B (LAT-B, 2) by Dr. Kashman and coworkers revealed an architecturally novel class of natural products.

The latrunculins are the first marine macrolides known to contain 16 (LAT-A)- or 14 (LAT-B)- membered rings (A). A tetrahydropyran (THP) ring links the macrolide to the thiazolidinone moiety (B). These compounds are outstanding in that they appear to be the first natural products to possess the thiazolidinone-2-one moiety which is rare in nature (C).



1 2

Numerous derivatives of LAT-A and LAT-B were synthesized in which each of the three parts of the molecule was modified separately. These were assayed for their biological activities. The relationship between chemical configuration and biological activity that emerged from these studies can be summarized as follows: 1) The macrolide is essential for permeation through the membrane. 2) The diene moiety in LAT-A protects the compound from degradation by serum. 3) The integrity of the THP ring, as well as the aminic hydrogen, are required for the specific effects of latrunculins, suggesting that latrunculin binding to its binding site occurs via hydrogen bonding, 4) Derivatives in which the THP ring was opened and derivatives in which the aminic hydrogen was modified, with the exception of N-hydroxymethylene-latrunculins, were cytotoxic but their effects differed from those of the native toxins.

The results of in vitro experiments on purified actin, carried out in collaboration with Dr. Korn at the NIH, showed that LAT-A acts directly on actin and that it affects the polymerization of pure actin in a manner consistent with the formation of a nonpolymerizable 1:1 molar complex with G-actin. These effects are qualitatively indistinguishable from the properties of G-actin binding proteins such as profilin and are very different from those of cytochalasins which are known to bind to the rapidly growing (barbed) end of F-actin filaments and to block all association.

In a recent series of experiments Dr. Spector and colleagues compared the short- and long-term effects of latrunculins on cell shape, actin organization, and cell growth and division to those of cytochalasin D, one of the most potent and specific members of the cytochalasin family. The results of the short-term comparison showed that LAT-A and LAT-B are more potent than cytochalsin D and revealed striking differences in the effects of the two classes of drugs on cell shape and actin organization. They also showed that the concentration-effect relationship for LAT-A in vivo reflects remarkably

well the dissociation constant (Kd) for the interaction of LAT-A with actin in vitro. This was not the case for cytochalasin D. All three drugs inhibited cytokinesis in synchronized cells, but their long-term effects were markedly different. Fibroblast cells treated with LAT-A stopped growing and maintained their altered state for extended periods. In contrast, the effects of cytochalasin D became cytotoxic with time in culture and the LAT-B-induced changes were transient in the continued presence of the drug. These transient effects were found to be due to a gradual inactivation of LAT-B by serum and were used to compare recovery patterns of cell growth, shape, and actin organization in two different cell types. This comparison showed that the transient effects of LAT-B were fully reversible for a normal fibroblast cell line and not for a transformed neuronal cell line. If, indeed, the biochemical property of latrunculins demonstrable in vitro does account for their dramatic effects in vivo, then small changes in the concentration of cellular actin may have far-reaching implications for cell morphology and behavior, and the regulatory roles of actin monomer-binding proteins such as profilin, as well as the regulation of actin synthesis, may be even more important than previously recognized.

"Biochemical Mechanisms in Toxic Cell Death"
R01 GM 33479 (Shier, W.), University of Minnesota

Elevated intracellular Ca^{++} appears to play a central role in the biochemical mechanisms of cell death occurring in various physiological and pathological situations. In these situations some mechanism causes the collapse of Ca^{++} gradients normally maintained by cells. For example, in ischemic diseases energy deprivation is believed to lead to collapse of Ca^{++} gradients by preventing the cell from continuing to pump out Ca^{++} ions. In some forms of toxic cell death membrane permeabilization may result in collapse of ion gradients. The resulting elevated intracellular Ca^{++} appears to trigger a biochemical pathway that results in general membrane permeabilization and cell death. Dr. Shier's laboratory is attempting to elucidate this pathway in a model system consisting of cultured fibroblasts treated with divalent cation ionophores (e.g., A23187) to elevate intracellular Ca^{++} and using trypan blue dye uptake to monitor membrane permeabilization.

By removing or reducing the concentrations of various ions in the culture medium during cell killing with Ca^{++} and A23187, it has been possible to define at least three steps in the membrane permeabilization process. The three steps are dependent on 1) low ($\geq$ 25 μM) extracellular Ca^{++}, 2) physiological extracellular Ca^{++}(1-2 mM), and 3) near physiological extracellular Na^{+} ($\geq$ 50 mM), respectively. The results obtained suggest that each of these are multistep processes with the first step being the most complex. It is inhibited by Mn^{++} and reversible unless the cell is stressed in one of a variety of ways (e.g., osmotic, oxidative, or high calcium stress). Ca^{++}-induced arachidonic acid release is also associated with the first step. The arachidonic acid does not appear to play any direct role in the membrane permeabilization mechanism. However, it may serve to monitor production of another metabolite that is a mediator. Among the plausible candidates are inositolphosphates. Consistent with this suggestion, this team finds that the amounts of phosphatidylinositol with diphosphate in cell membranes are reduced during Ca^{++}-induced membrane permeabilization. The amounts of the major structural lipids of mammalian membranes do not change

substantially during Ca^{++}-induced membrane permeabilization, but significant increases in amounts of lysophosphatidylcholine and lysophosphatidylinositol were observed.

Variant cell lines selected for resistance to A23187 plus Ca^{++} survive treatment with about 10-fold higher concentrations of ionophore than killed wild-type cells. Several different resistance mechanisms appear to be employed by various resistant clones. The most common mechanism appears to be synthesis of P-glycoprotein which may protect cells by transporting A23187-Ca^{++} complexes out of the cell before they dissociate.

The long-term goal of the work is the development of drugs that would protect vital tissues from Ca^{++}-mediated killing by inhibiting the membrane-permeabilization pathway until conditions improve. Key to the development of this type of drug is a complete understanding of the biochemical pathways of cell killing.

"Regulation of Calmodulin Dependent Processes"
R01 GM 34334 (Weiss, B.), Medical College of Pennsylvania

Calmodulin is a multifunctional calcium receptor in all eukaryotic cells. It affects a variety of cellular processes, including neurotransmission, muscle contraction, and cellular growth and differentiation. Early work from Dr. Weiss' laboratory showed that calmodulin can be inhibited by a number of pharmacological agents including neuropeptides. More recently, Dr. Weiss has developed and studied novel agents which are more selective in inhibiting the activity of calmodulin, for they act directly on the calmodulin-sensitive enzymes rather than on calmodulin itself. Two such compounds were synthesized. One is a chlorpromazine-calmodulin adduct (CPZ-CaM) and the other is a fluphenazine-N-mustard-calmodulin adduct (FNMCaM). These compounds were synthesized, purified, chemically characterized, and examined as to their effects on a number of calmodulin-sensitive and calmodulin-independent enzymes. The results showed that both CPZ-CaM and FNM-CaM were more potent than the free phenothiazines in inhibiting the CaM-sensitive phosphodiesterase. In addition, CPZ-CaM showed a great deal of selectivity in inhibiting the calmodulin-sensitive phosphodiesterase (IC50 = 0.2 uM) when compared with its ability to inhibit calmodulin-dependent ATPase (IC50 > 3 uM). The FNM-CaM adduct failed to show this selectivity. Although both free phenothiazines inhibited the activity of protein kinase C, neither of the phenothiazine-calmodulin adducts had any substantial effect on this enzyme. There were also differential inhibitory effects of the adducts on myosin light chain kinase. These data show that phenothiazine-calmodulin adducts can selectively inhibit certain calmodulin-sensitive enzymes and may be useful to explore further the varied biological effects of these enzymes.

Other studies in Dr. Weiss' laboratory focused on the regulation of calmodulin at the level of gene expression, using *in situ* hybridization histochemical techniques. They have designed two oligonucleotide probes for calmodulin mRNA complementary to specific regions of rat mRNA which code for either the Ca++ binding domain II or the N-terminal portion of calmodulin. These probes were characterized extensively to confirm their specificity of hybridization to calmodulin mRNA. The results of emulsion autoradiographic studies of rat brain sections showed that the probe bound to neurons rather than to glia or

white matter. Using in situ hybridization techniques this group found a heterogeneous distribution of calmodulin mRNA in rat brain, corresponding to the distribution of calmodulin as measured by immunocytochemistry or by the biological activity of calmodulin. Studies of the ontogenesis of calmodulin mRNA indicated that the message increased at different rates in the different brain areas. Future studies will be directed at uncovering the factors controlling the activity of calmodulin mRNA with the goal of developing selective pharmacological agents which will alter the production of calmodulin in specific cell types.

"Colon-Specific Drug Delivery"
RO1 GM 35147 (Friend, D.), SRI International

The overall goal of this research is to develop improved delivery systems for the treatment of inflammatory bowel disease (IBD). The delivery system is designed to release active drug in the colon (large intestine) so that relatively high concentrations of drug can be obtained in the inflamed tissues. The drugs being investigated are antiinflammatory steroids, such as dexamethasone and prednisolone and the nonsteroidal antiinflammatory agents 4- and 5-amino salicylic acid (ASA). The approach used to deliver these drugs to the large intestine is through transient derivatization of the drug to form a prodrug. These prodrugs, due to increased size and hydrophilicity, are poorly absorbed from the gastrointestinal tract. The prodrugs (glycosides if steroids and hemisuccinate or hemiphthalates if ASA compounds) are hydrolyzed in the large intestine by enzymes produced by gut microflora that are found in very large numbers in the colon and cecum. The animal model currently under investigation is the guinea pig.

A number of in vitro and in vivo experiments have been performed to evaluate the efficiency of drug delivery to the large intestine of the guinea pig. Most of these studies have been performed in the laboratory of Dr. Thomas N. Tozer. Transit time for delivery of the prodrug forms following gastric intubation is 1 to 2 hours in the fasted guinea pig. In fasted rats the transit time is closer to 5 hours. There is very little hydrolysis of prodrug in the stomach or small intestine, however, the prodrugs are very rapidly hydrolyzed once the prodrug enters the large intestine. A number of in vitro experiments help explain the results observed in vivo. Rates of hydrolysis of dexamethasone-glucoside in vitro indicated that glucosidase activity was at least 10-fold greater in the cecum and colon than the stomach, distal small intestine, and proximal small intestine. The half-life of dexamethasone-glucoside in cecal contents was about 10 minutes for unhomogenized contents and about 15 minutes for homogenized contents. A very interesting finding is that the guinea pig has unexpectedly high levels of β-glucosidase activity in the small intestinal tissues. The half-life of dexamethasone glucoside in tissues homogenates from the stomach, proximal small intestine, distal small intestine, cecal, and colon tissues were 133 minutes, 7 minutes, 25 minutes, 90 minutes, and 86 minutes, respectively.

Current work is focused on measuring the efficacy of the delivery system in vivo in guinea pigs in which inflammatory bowel disease has been induced and the measurement of the pharmacokinetics of drug delivery to the lower intestine in normal and IBD model guinea pigs.

"Effect of Protein Binding on Renal Drug Transport"
R01 GM 35498 (Smith, D.), University of Michigan

The kidney is a multifunctional organ that serves as a site of drug action, a target of drug toxicity, a focus of drug therapy, and the primary route of elimination for a number of drugs and metabolites. Dr. Smith and coworkers have studied three different aspects of renal drug transport: the effects of protein binding on renal extraction, the effect of altered blood flow on clearance, and the effects of blood flow on renal tubular excretion. Dr. Smith has been refining a mathematical description of biological determinants of renal drug clearance.

The role of protein binding on renal drug secretion is vague and poorly understood. The first series of studies evaluated in a systematic manner the effect of protein binding on the renal extraction of chlorothiazide in the isolated perfused rat kidney. Chlorothiazide was introduced into the recirculating perfusate and was assayed using high-performance liquid chromatography. Functional viability of the kidney was assured by the fractional excretion of sodium and glucose and by glomerular filtration rate. The protein binding of drug in perfusate was determined by equilibrium dialysis. These experimental conditions resulted in an approximate 14-fold increase of percent free drug (from 5.3 to 72.0 percent) and a three-fold increase of renal clearance and secretion. The data were best fitted to a model in which one Michaelis-Menten term was used to describe active transport, and secretion was dependent upon free concentrations of chlorothiazide in the perfusate. The maximum velocity of secretion (Vm = 85.6 μg/min) and Michaelis constant (Km = 2.1 μg/ml) of chlorothiazide were estimated using a nonlinear least-squares regression program. These results suggest that for compounds of low renal extraction, free drug concentrations are the driving force for carrier-mediated tubular secretion.

Studies evaluating the effect of changes in renal blood flow on drug excretion via the kidney have been almost nonexistent. Consequently, the role of blood flow on renal drug extraction is still considered in a descriptive rather than quantitative manner. In a second series of studies, angiotensin II was used as a probe to study the effect of changes in perfusate flow rate on the renal clearance parameters of furosemide in the isolated perfused rat kidney. Angiotensin II was found to have a dramatic effect on the renal hemodynamics, resulting in a 42 percent decrease in perfusate flow, a 27 percent decrease in glomerular filtration rate (GFR), and a 25 percent increase in filtration fraction. Values for the fractional excretion of glucose were very low and consistent with or without angiotensin II. Although the fraction of unbound furosemide was unchanged between treatments, the renal and secretion clearances of furosemide were reduced by about 30 percent in the presence of angiotensin II. However, if the renal clearance (CL_r) was corrected for free fraction (fu) and GFR [$ER=CL_r/(fu.GFR)$], there was no difference between the excretion ratio (ER) of furosemide in both treatments. These results imply that the altered clearance parameters of furosemide are more likely the consequence of a reduction in functional nephron mass rather than a change in intrinsic secretory transport per unit mass of nephron.

In observing the efficiency of renal extraction of para-aminohippurate (PAH), the prototype organic anion used to study renal secretion, in the isolated

perfused rat kidney (rat IPK), Dr. Smith and coworkers noticed that a significant portion of the PAH dose was being eliminated as metabolite. In a final series of experiments, the renal transport and metabolic fate of PAH in the rat IPK were determined. The total clearance of PAH (CL_T) was comprised of the renal clearance of unchanged PAH (CL_T) plus the N-acetyl-PAH that was formed by the kidney (CL_m). It was found that the N-acetyl-metabolite accounted for approximately 25 percent of the total elimination of PAH by the kidney, regardless of whether BSA was present or absent in the perfusate. If one were not aware of this metabolite in the rat IPK, a substantial underestimation of the efficiency of PAH excretion and extraction would be made. This data is also consistent with in vivo studies in the rat as well as in patients with renal impairment.

"Adenylyl Cyclase and the Plasma Membrane"
RO1 GM 36259 (Neer, E.), Bringham and Women's Hospital

The G proteins are a family of heterotrimeric proteins which play a central role in transducing hormonal and sensory signals across the plasma membrane. They are, therefore, a critical point of control for a variety of agents which modulate cellular functions. Over the past 5 years, work from Dr. Neer's and other laboratories has revealed an unexpected diversity of isoforms of G protein α, β, and γ subunits. Recent work from Dr. Neer's laboratory has focused on studies of the structure and function of the G protein. The laboratory has taken three approaches to the problem: 1) identification of functional domains by chemical modification and generation of active proteolytic fragments, 2) cloning and characterization of cDNA's for new α subunit subtypes, and 3) reconstitution of α subunits with membrane effectors.

Dr. Neer's laboratory has characterized the reactive sulfhydryl groups on the major G protein α subunit from brain, α_o. The laboratory found that in unliganded α_o, there are three sulfhydryls that are reactive, but only one when the ligand GTPγS is bound. The single sulfhydryl, which is reactive in the liganded state, was found to be located at position 108 in the variable region of the α subunit sequence. Reaction of this sulfhydryl group with the alkylating agent, N-ethylmaleimide, changes the conformation of α_o, since it prevents modification of the protein by pertussis toxin at a distant site. Analysis of α subunit structure by limited proteolysis has shown that an amino terminal 2 kD peptide can be removed by proteolysis with trypsin, and the resultant truncated molecule cannot interact with $\beta\gamma$. Cleavage of the carboxyl terminal 1 or 2 amino acids with carboxypeptidase does not affect the ability of the α subunit to associate with $\beta\gamma$, but does entirely eliminate the covalent modification by pertussis toxin, probably by destroying the substrate site for that enzyme.

Analysis of purified proteins in Dr. Neer's and other laboratories had suggested the existence of multiple α subunits which could be substrates for covalent modification by pertussis toxin. Using molecular cloning techniques, Dr. Neer's laboratory isolated a cDNA clone from a bovine pituitary library which encoded an, as yet, unidentified α subunit. Using the bovine cDNA's as probes, Dr. Neer's laboratory analyzed a human T cell library and identified a third type of α subunit. In collaboration with the laboratories of Dr. Jonathan Seidman and Dr. Thomas Shows, Dr. Neer showed that the genes for 2 α subtypes were found on different human chromosomes.

A major interest in Dr. Neer's laboratory has been to define the function of α and $\beta\gamma$ subunits. A collaborative study with Dr. David Clapham and his colleagues showed that $\beta\gamma$ subunits from bovine brain can activate the cardiac muscarinic K^+ channel. In further studies, Drs. Neer and Clapham showed that activation of the K^+ channel could also be mediated by the α subunits. The effects of α and $\beta\gamma$ are not additive, so that both subunits seem to affect the same population of channels. Activation of the muscarinic K^+ channel by $\beta\gamma$ subunits seems not to be a direct interaction of the $\beta\gamma$ subunits with the channel, but instead to be mediated by activation of phospholipase A_2. Activation of the channel by both α and $\beta\gamma$ subunits implies a more complicated scheme for G protein regulation of the channels than had previously been supposed.

Correct transduction of signals across the plasma membrane is essential to maintain normal cell growth and function. Studies over the past 2 years have begun to define the structural basis for the function of some hormone-responsive membrane proteins. Such information may, in the future, allow development of drugs that modulate hormonal responses, and perhaps lead to ways of correcting the abnormal function of this system in diseased or malignant cells.

"Chemical Detoxication During Hypoxia"
RO1 GM 36538 (Jones, D.), Emory University

Hypoxia is a common clinical variable that can affect the therapeutic efficacy, toxicity, and elimination of pharmacological agents. During the past year, Dr. Jones has found that the fundamental changes in cell function during oxygen deficiency result in a general increase in susceptibility of cells to chemical-induced injury. This is especially evident in the glutathione-dependent pathways where systems that normally maintain the pool by biosynthesis, reduction of the disulfide form, or Na^+-dependent uptake are all impaired due to hypoxia. In addition to these changes, there is a suppression of energy production that is maintained for up to 2 hours after short-term anoxia, indicating that there is a general suppression of energy-dependent metabolism and repair functions during post-anoxic recovery. Cellular ionic balances of sodium, potassium, and calcium are similarly perturbed. Thus, the cells are less able to detoxify exogenous compounds and are more vulnerable to injury.

Other studies showed that these effects are further amplified by chronic exposure to hypoxia. Cells from animals adapted to hypoxia for 2 weeks were found to be paradoxically more sensitive to hypoxic injury. The activities of several of the major drug detoxication enzymes are decreased by this adaptation, and Western blot analyses indicate that the expression of the proteins is decreased due to hypoxia. Because other studies have shown that hypoxia induces an increased expression of many detoxication enzymes, the current studies indicate that expression could be dynamically regulated over the entire physiological oxygen concentration range. Studies are ongoing to carefully define the time courses of these changes with the intent to explore the mechanisms involved in oxygen-dependent control of expression of detoxication systems during hypoxia.

Although many drugs appear safe when administered to hypoxic individuals, the recent studies show that hypoxia can result in prolonged plasma half lives, delayed intestinal absorption and renal elimination, decreased metabolism, and increased toxicity. Future studies will address the detailed mechanisms involved in these processes and focus on the effects of hypoxia on therapeutic agents that are frequently used in hypoxic individuals.

"Catatoxic Steroids: Molecular Mechanism of Action"
R01 GM 37044 (Parkinson, A.), University of Kansas Medical Center

One of the aims of this team's research is to determine how treatment with certain steroidal agents (known as catatoxic steroids) can protect rats from the adverse effects of numerous chemicals, such as drugs, pesticides, and toxicants. The working hypothesis is that catatoxic steroids (e.g., dexamethasone and pregnenolone-16oc-carbonitrile [PCN]) induce a form of liver microsomal cytochrome P-450 that metabolizes and detoxifies a wide variety of foreign chemicals. A steroid-inducible form of cytochrome P-450, designated cytochrome P-450p, has been purified and an antibody prepared in rabbits immunized with this rat liver protein. The antibody is highly specific for cytochrome P-450p, and does not cross-react with any of the other known forms of rat liver cytochrome P-450.

In contrast to all other forms of cytochrome P-450 purified in this laboratory, cytochrome P-450p was catalytically inactive when solubilized and purified from liver microsomes. Therefore, to determine the catalytic activity of cytochrome P-450p, Dr. Parkinson and coworkers measured the metabolism of selected chemicals by rat liver microsomes in the presence and absence of antibody against cytochrome P-450p. These antibody-inhibition experiments established that cytochrome P-450p plays a major role in the metabolism of certain drugs (such as digitoxin), steroids (namely testosterone), and chemical carcinogens (namely aflatoxin B_1 and benzo[a]pyrene). These results support the hypothesis that the protective effect of catatoxic steroids is related to their ability to induce a form of liver microsomal cytochrome P-450 that can metabolize and detoxify a wide variety of chemicals.

The antibody-inhibition experiments described above were also performed with liver microsomes from mice, hamsters, guinea pigs, rabbits, cats, dogs, cynomolgus monkeys, and humans. In all cases, cytochrome P-450p was shown to play a major role in the metabolism of testosterone, which, with the exception of the rabbit, was converted by cytochrome P-450p to 2β-, 6β-, and 15β-hydroxytestosterone. However, in species other than the rat, cytochrome P-450p played little or no role in the metabolism of digitoxin. The rat was also unique in that it converted digitoxin to digitoxigenin bisdigitoxoside, whereas the other species tested did not. These results were initially interpreted to mean that cytochrome P-450p has different functions in different species. However, further studies have provided an alternative explanation for the apparent species-dependent function of cytochrome P-450p.

The alternative explanation stemmed from the finding that rat liver microsomes contain two electrophoretically distinct forms of cytochrome P-450p (as determined by Western immunoblot). The higher molecular weight form of

cytochrome P-450p (MW 51,500) appears to be responsible for metabolizing testosterone, whereas the lower molecular weight form (MW 51,000) appears to be responsible for metabolizing digitoxin. It is possible, therefore, that the observed species differences in the functions of cytochrome P-450p arise because the higher molecular weight form of cytochrome P-450p is conserved from species to species, whereas the lower molecular weight form is specific to the rat. Dr. Parkinson and coworkers are currently testing this possibility, and they are attempting to determine which form of cytochrome P-450p is responsible for metabolizing the carcinogens, aflatoxin and benzo[a]pyrene. The results of these ongoing experiments could have important consequences for understanding and predicting species differences in the metabolism and toxicity of foreign chemicals.

"The Characterization of Human Liver Cytochromes P-450"
R01 GM 37498 (Wrighton, S.), Medical College of Wisconsin

One of the most abundant forms of human liver cytochrome P-450 found in the liver is the steroid-, macrolide antibiotic- and phenobarbital-inducible cytochrome P450HLp, termed HLp. Studies using recombinant DNA techniques indicate that HLp is only one member of a family of related proteins which contains up to five members. However, the actual expression of the HLp-related forms has only recently been demonstrated by Dr. Wrighton's group. Using techniques that identify immunochemically related proteins, Dr. Wrighton's group found that one out of five adults express, in addition to HLp, a second highly related form. In addition, approximately 50 percent of the total cytochrome P-450 present in the human fetal liver is accounted for by a single form of cytochrome P-450 that is related but not identical to HLp.

Dr. Wrighton's group purified these two HLp-related proteins and compared them by a number of biochemical and functional criteria to HLp. The results of these comparisons revealed that the fetal form of cytochrome P-450 (referred to as HLp2) and the new adult form (referred to as HLp3) are indeed highly structurally related to HLp but yet distinct gene products. In addition, several substrates metabolized by HLp are also metabolized by HLp3. Thus this family of human liver cytochromes P-450 contains at least three separate gene products.

These findings are of great significance with respect to human drug and carcinogen metabolism. First, HLp has been shown to exclusively metabolize the antihypertensive agent nifedipine and a polymorphism in the metabolism of nifedipine in one out of five people has been reported. Thus the expression of HLp3 in only one out of five patients may be related to the problem in the metabolism of nifedipine. This relationship is currently under investigation. The importance of finding a large amount of HLp2 in fetal livers is that the rat fetal liver does not contain a protein related to HLp2 and therefore studies using the rat as a model for human fetal drug and carcinogen metabolism should be interpreted with caution.

"Endogenous Modulation of Oto- and Nephrotoxicity"
R01 GM 40858 (Hoffman, D.), Dartmouth Medical School

The thiol-containing tripeptide glutathione (GSH) is part of an endogenous mechanism which protects against toxic insults caused both by reactive drug metabolites and by toxic forms of oxygen formed during intermediary metabolism. The hypothesis on which this research is based is that these mechanisms of protection mediated by GSH play an important role in the oto- and nephrotoxicity of loop diuretics, aminoglycosides, cisplatin, and similar compounds. This hypothesis offers the possibility of ameliorating the toxicities of many of these drugs by modifying their metabolism.

Major findings to date indicate that depletion of GSH by buthionine sulfoximine treatment or by semistarvation strongly enhances the toxicities of loop diuretics, aminoglycoside antibiotics, and cisplatin. Dr. Hoffman has found that cisplatin treatment directly depletes cochlear and kidney GSH levels. In animals pretreated with buthionine sulfoximine to deplete GSH, which is not known to have any toxic potential, administration of this otherwise nonfatal dose of cisplatin results in 100 percent mortality. This further serves to underscore the potential significance of GSH protective mechanisms in cisplatin toxicity. On the other hand, the immunosuppressive drug cyclosporine A has demonstrated an unusual effect on GSH levels in these studies. While liver GSH content appears to decline after cyclosporine administration, GSH levels in kidney actually increase. This increase in kidney tissue is quite marked and may reflect underlying mechanisms of toxic damage.

Another new finding of this work is that ethacrynic acid inhibits GSH reductase. This inhibition occurs in the same concentration range as does the well-known inhibition by ethacrynic acid of GSH S-transferase. The action of ethacrynic acid on the S-transferase has been regarded as a potential mechanism for toxicity of this drug, and for its toxic synergy with other drugs. The finding that ethacrynic acid apparently is equally active in inhibiting the recycling of oxidized GSH (GSSG) back to the biologically active reducing agent GSH opens up a new area of research in the toxic mechanisms of this drug. This mechanism may enhance toxicity of other compounds which generate peroxides or other active oxygen species. This enhancement of toxicity focuses attention on pathways other than the metabolite-conjugating pathway mediated by GSH S-transferase. Cis- and transplatin, aminoglycoside antibiotics, buthionine sulfoximine, furosemide, bumetanide, and piretanide were without effect on GSH reductase activity. Of this list, only the loop diuretics were found to inhibit GSH S-transferase. A further point of interest in this regard is that ethacrynic acid is approximately equivalent to BCNU (carmustine), a clinically used antineoplastic agent, in its ability to inhibit GSH reductase. This finding may have implications for the use of ethacrynic acid to enhance the therapeutic activity of antineoplastic agents, and may shed light on the mechanism of BCNU activity and toxicity.

Localization studies of GSH S-transferase and cytochrome P-450 enzyme immunoreactivities also have been performed. Using antisera against GSH S-transferase donated by Dr. Yogesh Aswasti, immunoreactivity was found in the stria vascularis of the inner ear. This tissue is known to be very

metabolically active, and to be responsible for much or all of the ion transport which occurs in the inner ear to maintain the ionic characteristics of the endolymph and perilymph. This enzyme, and its localization in tissues with possible functional similarities with kidney, may help explain the common toxicities of many drugs for kidney and ear, and may offer important data toward preventing the toxicities which limit the usefulness of these drugs.

ANESTHESIOLOGY RESEARCH OVERVIEW

The principal programmatic initiative in anesthesiology over the past 2 years has been advanced through a program announcement on the molecular pharmacology of anesthetic action, which appeared in the **NIH Guide for Grants and Contracts** in January 1988. The program announcement emphasized the Pharmacological Sciences Program's renewed interest in multidisciplinary research proposals aimed at clarifying at the molecular level the mechanisms of action of general and local anesthetic agents.

Proposals on a variety of topics have been and continue to be submitted in response to the program announcement. Among the proposals that have been funded to date are studies on: the molecular pharmacology of genetic sensitivity to volatile anesthetics; the action of volatile anesthetics on synaptic receptors; the electrophysiological effects of general anesthetics on isolated membrane patches; the molecular actions of local anesthetics on peptides isolated from ion channels; the effects of anesthetics on opiate peptide gene expression in spinal cord neurons; and the molecular target for steroid anesthetics.

The anesthesiology research grant portfolio has expanded by approximately 15 percent over the past year, in large measure due to an increased number of applications submitted in response to the program announcement. Importantly, this expansion has occurred without any decrease in the number of anesthesiology research grants that have traditionally comprised the anesthesiology grant portfolio, or a decrease in other program areas.

The research grant portfolio in anesthesiology covers the same basic areas supported through the pharmacology portfolio. These fields of study include: effects of general and local anesthetic agents on integrated organ systems; development of new delivery systems for anesthetics; distribution of anesthetics in the whole animal and in specific organs; metabolism of anesthetics; molecular pharmacology of anesthetics; interactions of anesthetics with other therapeutic agents; and toxicity of anesthetics, with a special emphasis on malignant hyperthermia and halothane-induced hepatitis.

Possibly the most exciting research area in anesthesiology supported by NIGMS since the last biennial report involves studies on possible specific sites of anesthetic action. A number of investigators supported by the Pharmacological Sciences Program are currently working on this general problem. A few key advances are highlighted below.

Two years ago, PS Program grantees were able to identify different types of binding of halogenated hydrocarbon, volatile anesthetics in brain by using nuclear magnetic resonance spectroscopy. The data indicated that there may be tight and saturable binding of these agents to specific regions of the brain. Further evidence for specific, saturable sites of anesthetic binding has continued to be reported, although the actual sites have yet to be identified.

The possibility is being explored that the G proteins that are linked to a variety of cell surface receptor proteins may be targets for anesthetic action. While evidence for this hypothesis is not yet conclusive, it certainly is suggestive of the possibility that there are specific sites of anesthetic action in neuronal tissues.

The fascinating discovery has been made that α_2-adrenoreceptors in brain may play an important role in the maintenance of the state of general anesthesia. This work has expanded on research previously supported by the PS Program on the role of α receptors in the modulation of pain pathways in the spinal cord. Although it is not yet clear whether the α_2-adrenoreceptor phenomenon applies to all volatile anesthetics, it does add to the increasing body of evidence suggesting specific sites of anesthetic action in brain.

Biophysical approaches are being used to study anesthetic action in the belief that the effects of anesthetics might still be described most accurately by carefully delineating the interactions of anesthetic agents on lipid and protein components of the cell membrane. Many of these studies employ artificial membranes in which the protein and lipid content can be carefully controlled. Unique effects of halogenated hydrocarbons with a strong dipole moment on the interaction of interfacial water molecules with phosphate components of the membrane have been reported. Studies have been clarifying the direct and indirect effects anesthetic agents may have on membrane-associated proteins such as the nicotinic receptor and phospholipases when these compounds dissolve in sections of the cell membrane adjacent to the protein of interest. These studies involve complex kinetic analyses of anesthetic interactions with components of the membrane and may eventually produce a fusion of the contrasting hypotheses that expound either specific or non-specific effects of anesthetics.

The Pharmacological Sciences Program looks forward to a continuing growth in the types and numbers of research grants in anesthesiology. With investigators from a variety of disciplines becoming interested in research problems germane to anesthesiology and with a larger pool of M.D.'s being trained to pursue research careers in anesthesiology, the upcoming years promise to be particularly fruitful for anesthesiology research. Importantly, this is an area where research findings have quick benefit in improving clinical care.

"Anesthesiology Center for Research and Teaching"
P50 GM 15904 (Miller, K.), Harvard Medical School

This anesthesiology research center is devoted toward elucidating the molecular mechanism of action of anesthetic agents. The projects within the center are focused on clarifying the interaction of anesthetics with a variety of cellular and subcellular components. Multidisciplinary approaches are

being used to explain how general anesthetic agents affect model receptor systems and how local anesthetics bind to various components of the cell membrane.

- **Mechanism of Action of General Anesthetics and Pressure of Cholinergic Receptors** (Miller, K.)

General anesthetics are thought to act at synapses in the central nervous system. Although the physiological site of action remains enigmatic, the molecular nature of this site of action is known to be relatively hydrophobic and not to be sensitive to the structural geometry of the general anesthetic molecules. However, the nature of the site is well modeled by both bulk phases like lipid bilayers and by a site on a pure protein such as the enzyme luciferase. To resolve this problem Dr. Miller and his colleagues have been studying an intact postsynaptic membrane, the acetylcholine receptor-rich membrane from Torpedo electroplaque.

General anesthetics have several actions on this system and the overall pharmacology is quite complex. However, their most potent action is to inhibit ion channels opened by acetylcholine. Recently, Dr. Miller has exploited the newly discovered self-inhibitory site, a site that enables high concentrations of acetylcholine to "block" its own channel, to approach the question of whether general anesthetics act at a discrete recognition site.

The agonist concentration-response curve for ion flux from acetylcholine receptor-rich membrane vesicles is bell shaped, the rising phase representing the action of two agonist molecules in opening each channel and the falling phase representing the action of channel inhibition at higher concentrations. If anesthetics act at the same inhibitory site as agonists, then they should compete with the agonist for this site. Dr. Miller's data show that procaine, a typical channel-blocking local anesthetic, acts this way. When channels are heavily inhibited with procaine the falling phase of the agonist curve is shifted to the right consistent with competitive interaction between the two inhibitors at a common site. Procaine and acetylcholine share some structural features, so this result might be anticipated. It was therefore of interest to determine if general anesthetics act at the same site.

Dr. Miller's results showed that octanol and procaine do not compete for a common inhibitory site, demonstrating that general and local anesthetic action differ even though kinetically they share some common features. In fact, the inhibition curve for octanol alone does suggest a single site of action, raising the possibility of one site of action for procaine and one for general anesthetics. However, when the interaction between octanol and hexanol was studied, it was found that these close homologs did not compete for a single inhibitory site. Thus, either each alcohol has its own site or, more probably, some other mechanism is at work.

One such mechanism might be the interaction of the alcohols with a large hydrophobic patch on the protein. The acetylcholine receptor is a large integral membrane protein dependent on lipid for its action. The interface between the protein and its surrounding lipid bilayer provides one possible target for anesthetic action. Dr. Miller and his coworkers are now studying the effect of general anesthetics at this interface; preliminary data show

that general, but not local, anesthetics displace phospholipids from the lipid-protein interface.

"Interfacial Action of Inhalation Anesthetics"
RO1 GM 25716 (Ueda, I.), University of Utah

Anesthesia is generally attributed to an increase in membrane fluidity. Because the membrane fluidity is a property of the conformation of the tails of lipid molecules, the effect of anesthetics might be attributed to their direct action on the core of the membrane. Nevertheless, all clinical volatile anesthetics are dipolar molecules containing acidic protons. These molecules are amphipathic and expected to stay at the membrane-water interface. Apolar molecules, such as alkanes, that penetrate into the membrane core are weaker anesthetics despite their direct interaction with the lipid tails. This apparent paradox between surface binding and core action has been addressed in this project.

Experimental evidence of interfacial binding of polar anesthetics is provided by two-dimensional proton nuclear Overhauser spectrometry. A cross-peak was observed between the hydrophobic bond of methoxyflurane and hydrophilic choline moiety of DPPC (dipalmitoylphosphatidylcholine) membrane. No other cross-peaks were found. This anesthetic adsorbs at the membrane-water interface, interacting only with the hydrophilic surface of the membrane. The hydrophilic domain of the anesthetic molecules stays in water.

Fluorine-NMR revealed that binding of halothane to micelles occurred in two distinctly different modes. At low concentrations, the binding followed the Langmuir adsorption isotherms to saturate the surface of the membrane. After saturation, the binding mode changed to multilayer stacking (no apparent saturation) following the condensation kinetics according to the Bose-Einstein statistics.

Dr. Ueda has proposed that the anesthesia mechanisms involve destruction of the hydrogen-bonds (H-bond) between the membrane and the water molecules clustered around the macromolecular surface. Lipid membranes are supported by the H-bonded matrix of water molecules. When the interaction force is weakened, membranes become disordered (fluidized) and expand, as has been speculated in previous studies.

The H-bond breaking activity of anesthetics was estimated by differential scanning calorimetry. The surface-bound water molecules are already in a crystalline form and freeze only at subzero temperature, as reported previously. The structure of water cluster at biological surface is strongly condensed because of the electrostatic force from the surface charges acting on the water dipole; a phenomenon known as electrostriction. The number of water molecules electrostricted at the membrane surface is determined by the surface charge density. This group has shown with Na-NMR and conductometry that volatile anesthetics increased the Na ion (counterion) binding to anionic surfactant micellar surfaces, thus decreasing the surface charge density, and released the electrostricted water.

The H-bond breaking site on the lipid membrane was estimated by Fourier-transform infrared spectroscopy in water-in-oil reversed micelles with DPPC-benzene mixture. Volatile anesthetics decreased the hydration of phosphate moiety and increased the free water, evidenced by the appearance of the free water O-H stretching and bending frequency. Despite the positive charge, the choline moiety did not contribute to the hydration, and did not show any change when volatile anesthetics were added. Studies with fully hydrated DPPC membrane are now in progress with Fourier-transform total reflection infrared spectroscopy.

"Anesthesia and Hypoxic Pulmonary Vasoconstriction"
RO1 GM 29628 (Marshall, B.), University of Pennsylvania

In a series of studies utilizing dogs, cats, rats, rabbits, pigs, sheep, and human subjects, the team of Dr. Bryan Marshall and Dr. Carol Marshall have demonstrated that hypoxic pulmonary vasoconstriction is a critical factor in determining changes in pulmonary hemodynamics and arterial oxygenation during general anesthesia.

The quantitative results of these studies, together with those from other laboratories, have been combined for the development of a generalized model of the pulmonary circulation. This computer model is based on the biodynamic properties of the vascular tree and now encompasses all of the common physiologic variables that may accompany anesthesia, including posture, mechanical ventilation, and thoracotomy, as well as changes in end expiratory pressure, cardiac output, left atrial filling pressure, and vascular tone at any generation of vessel. Furthermore, the model has been extended to multiple compartments so that the effects of $\dot{V}/\dot{Q}$ ratio distribution on hemodynamics and gas exchange can be evaluated.

Hitherto, the influence of anesthetic agents and techniques on the pulmonary circulation have been difficult to identify because the introduction of the common general anesthetic agents is inevitably accompanied by alteration of several variables simultaneously. The combination of experimental observation and analysis of principal variables with the computer model has proved to be a powerful tool for identifying both the dominant causes and the underlying physiologic mechanisms. These techniques have established the importance of inhibition of hypoxic pulmonary vasoconstriction as a cause of hypoxemia during general anesthesia in all subjects, as well as during thoracic surgery. Examination of a variety of drugs including almitrine, $PGF_{2\alpha}$, dopamine, dobutamine, serotonin, and histamine has revealed a bimodal dose-response effect on hypoxic pulmonary vasoconstriction. These results have revealed an underlying unity for much published work, as well as emphasizing the considerations necessary for achieving therapeutic effectiveness in clinical practice.

"Anesthetic Reactions in Surgery"
RO1 GM 31382 (Louis, C.), University of Minnesota

The long-term objective of this research has been to define the molecular basis of the defects that are responsible for the pharmacogenetic syndrome malignant hyperthermia (MH). In this anesthetic reaction that occurs in

susceptible individuals during surgery, certain anesthetic agents and depolarizing muscle relaxants trigger a rise in body temperature and muscle rigidity which can result in death. Using a purebred strain of malignant hyperthermia susceptible (MHS) pigs, Dr. Louis and coworkers have now demonstrated that the altered control of sarcoplasmic Ca^{++} in MHS muscle, likely results from specific defects in the skeletal muscle sarcoplasmic reticulum (SR) and transverse tubule (TT) membranes.

MHS SR demonstrated a significantly greater extent of Ca^{++}-induced Ca release than did normal SR. This correlated with the greater affinity of MHS SR for the plant alkaloid ryanodine. That this alkaloid binds to the open form of the SR Ca release channel supports the Ca release data indicating that the MHS SR Ca channel is more open, and can release more Ca than can the normal SR Ca channel. This conclusion was supported by single channel kinetic data indicating the MHS SR Ca release channel is not inactivated by Ca^{++} concentrations that inactivate the normal SR Ca release channel. Finally, initial studies indicate that the proteolytic sensitivity of the SR ryanodine receptor differs in MHS and normal muscle indicating a difference in the primary sequence of the MHS and normal ryanodine receptors.

MHS TT and normal TT were in all respects essentially identical except for the binding of calcium antagonists to the dihydropyridine receptor. Specifically, nitrendipine and desmethoxyverapamil binding to MHS TT was approximately 50 percent that of normal TT. This likely reflects an adaptation by the TT dihydropyridine receptor/voltage sensor to the primary defect in the SR Ca release channel which in MH is more easily triggered by activators of excitation-contraction coupling. Thus, a down-regulation of the dihydropyridine receptor/voltage sensor allows MHS individuals to lead a normal life, in spite of a defective SR Ca release channel.

Taken together, these data provide the first strong evidence that the SR Ca release channel/ryanodine receptor is the site of the primary defect in MH. This now provides a testable hypothesis as to the molecular basis of the syndrome.

"Axonal Study for Improved Selective Nerve Anesthesia"
RO1 GM 31710 (Fink, B.), University of Washington

Differences in the sensitivity of various nerve fibers to block by local anesthetics have been the target of much investigation. Dr. Fink's recent work has led to a new outlook on the problem. Small diameter fibers were long thought to be easier to block than large diameter ones. However, the series of studies by Drs. Fink and Cairns on individual mammalian nerve fibers have provided new evidence that the real differential factor is not the diameter of the fiber but the length of fiber exposed to anesthetic; in the case of a fiber insulated by myelin, the decisive factor is apparently the number of consecutive patches bare of myelin nodes of Ranvier that come into contact with the anesthetic. A large diameter fiber tends to have long distances between the nodes and, when a marginal concentration of anesthetic is used, a greater fiber length must be exposed to anesthetic to effect block than is the case with smaller diameter fibers, where the nodes are closer to each other. Moreover, the necessary length of fiber apparently increases with a weaker concentration of anesthetic. Independent workers in another laboratory,

besides confirming this, systematically demonstrated that the minimum bathed length of fiber required for block increases with the dilution of the anesthetic; they suggest that the anesthetic causes a cumulative decrease of electric current at successive bathed nodes, eventuating in block.

Dr. Fink has now presented a mathematical formula which summarizes his own and the independent results. Taking into account that at normal nodes the nerve impulse is ordinarily conducted with a large excess of electric current and that, by definition, conduction block occurs when the residual current or "safety of conduction" decreases to less than 1, his formula expresses the reduced current caused by local anesthetic as an inverse function of the product of the anesthetic potency, the anesthetic concentration, and the number of consecutive nodes bathed by anesthetic. In brief, the equation expresses in precise terms the minimum conditions for block of conduction by local anesthetic in a myelinated nerve fiber. This has altered the dimensions of the investigation, because the equation for the first time provides a quantitative frame for evaluating the differential block of modalities served by myelinated axons, such as pain, touch, and skeletal motor activity.

Dr. Fink has shown how his equation unifies numerous previously disparate and even seemingly contradictory data on differential block associated with clinical epidural and subarachnoid anesthesia. In particular, it clarifies the limited possibilities of differential block of pain by a local anesthetic. The equation also initiates an important extension to the pharmacokinetics of regional anesthesia. Previously, the pharmacokinetics have been concerned mainly with tracking the absorption of local anesthetic into the blood compartment and correlating this with the onset of toxic cerebral and cardiac effects. The study of regional distribution of local anesthetic can now usefully return to the compartments where the drug is injected and acts before it is absorbed into the blood stream. For example, in the case of subarachnoid anesthesia these compartments are 1) the spinal cerebrospinal fluid and 2) the roots of the spinal segmental nerves. By modeling the drug levels in each of these compartments in terms of the new equation, Dr. Fink can show how a pharmacokinetic calculation precisely predicts the dosages required for subarachnoid block by single, repetitive, or continuous injection.

"Epidural Clonidine Analgesia in Obstetrics"
R29 GM 35523 (Eisenach, J.), Wake Forest University Medical Center

Current methods of providing pain relief for women during labor have many bothersome or dangerous maternal and fetal side effects. Epidurally administered clonidine has been shown to produce segmental analgesia, without other sensory or motor blockade, by an α_2-adrenergic action in the spinal cord. The focus of these studies in sheep is to examine the analgesic utility and maternal and fetal side effects of epidurally administered clonidine.

In previous work supported by this grant, Dr. Eisenach and coworkers have demonstrated that epidurally administered clonidine produces dose-dependent analgesia in sheep by a local α_2-adrenergic mechanism. Clonidine does not alter neurobehavior, spinal cord histology, or spinal cord blood flow. When

injected i.v., clonidine produces transient hypoxemia. This hypoxemia is produced via a peripheral α_2-adrenergic mechanism and not accompanied by respiratory or cardiovascular depression. This effect does not occur in humans or dogs, and likely represents an unusual form of platelet aggregation in sheep.

Recently, this team has investigated the effects of epidurally administered clonidine in pregnant ewes, and the sites of action of its hemodynamic effects. Epidurally administered clonidine is rapidly absorbed into maternal plasma and extensively crosses the placenta, producing a decrease in fetal heart rate. However, maternal blood pressure, uterine perfusion, and maternal and fetal arterial blood gas tensions are unaffected by epidurally administered clonidine. In a study of the pharmacokinetics and pharmacodynamics of clonidine, these investigators have demonstrated that epidurally administered clonidine appears rapidly and has a high bioavailability in lumbar CSF and that the minimal effect of clonidine by this route on blood pressure represents a balance of depressor effects in the spinal cord and depressor effects in the peripheral vasculature. These findings have led to clinical studies of this therapy.

"Control of Neural Function in Anesthesia"
RO1 GM 35647 (Strichartz, G.), Brigham and Women's Hospital

When used in the current clinical methodology, local anesthetics (LA) block impulses in all sensory and motor axons of the peripheral nerve. The goal of this research has been to elucidate the features of certain functionally defined nerve fibers (e.g., "pain" fibers) that will increase their susceptibility to local anesthetics. To accomplish this Dr. Strichartz and coworkers have used excised peripheral (sciatic) nerves from frogs and rats and have also studied rat sciatic nerve in vivo, using a light general anesthesia (pentobarbital).

From the in vitro studies they have demonstrated that the concentration of LA required for impulse blockade (C_L) depends on the length of nerve exposed to the drug. Among relatively large myelinated fibers, C_L of lidocaine fell by 20-50 percent when the exposure length was increased from 6 to 21-26 mm. At one lidocaine concentration (0.8 mM) the compound action potential-amplitude fell by 90 percent when the exposure length was increased from 6 to 38 mm; however, much of this decrease is due to differential slowing of impulse in single fibers and not to actual extinction of the impulse. Of importance is the fact that no correlation was found between C_L and the impulse conduction velocity, usually taken as an indicator of axon diameter. This unequivocally disproves the widespread belief that C_L for small fibers is lower than that for larger fibers.

Studies in vivo have repeated and extended these findings. At one exposure length (2.0 cm, controlled by a perfusion chamber placed around the nerve in the rat's thigh), C_L ranged from 0.4 to 1.6 mM, for myelinated A-beta and A-delta fibers and non-myelinated C-fibers. The ranges of C_L for these different fiber classes overlap considerably, again showing a lack of correlation between size and LA susceptibility.

All the above findings used a single stimulus protocol and describe the response of a single impulse or volley. However, trains of impulses produced by repetitive stimulation reveal a differential response among classes of fibers. Impulses in non-myelinated C-fibers and small myelinated A-delta fibers are more susceptible to lidocaine at repetitive stimulation, often at frequencies <5-10 Hz. Large myelinated axons also show such "use-dependence," but to a smaller degree and at higher frequency. A separate phenomenon also occurs, in the absence of LA, where impulses in small fibers slow and fail "endogenously" during repetitive stimulation in the absence of drug. These are reversible processes which recover when the stimulation ceases or the drug is washed out.

Early evidence shows a striking difference between functionally identified classes of A-delta fibers. Those fibers shown to be nociceptors (fast pain) have impulses that "endogenously" slow greatly (doubling conduction times in the rats) and often fail with repetitive stimulation, whereas those shown to respond selectively to cooling have none of this behavior. Continuing studies of the drug sensitivity of such functionally identified axons will establish the prospects for exacting a truly modal selective blockade--the relief of pain without effect on other sensations or on motor activity.

"Action of Anesthetics on the Cardiovascular System"
R29 GM 38021 (Robinson-White, A.), Uniformed Services University of the Health Sciences

Anesthetics may interact in a non-specific manner with cellular membrane lipids and proteins or they may act specifically to alter the activity of membrane proteins. Studies on the effect of anesthetics on the cardiovascular system have, in the past, mainly focused on mechanical changes in the body as a whole, rather than on biochemical changes in cellular events.
Dr. Robinson-White and coworkers have recently revealed evidence of specific effects of barbiturates on a biochemical mechanism (phosphatidylinositol hydrolysis) in endothelial cells from large blood vessels of rat, and in an unrelated cell type, the rat basophilic leukemia (RBL-2H3) cell.

By use of a radiolabel assay for the products of the hydrolysis of inositol phospholipids, and a radioligand binding assay, Dr. Robinson-White has found evidence of specific changes in hydrolysis as affected by barbiturates. Barbiturates, at clinical and pharmacological concentrations, inhibited hydrolysis in endothelial cells by a mechanism other than competitive inhibition at the ligand-receptor; however, they inhibited ligand binding at the receptor site. Using chemically permeabilized cells, Dr. Robinson-White and her associates have found that barbiturates may inhibit both hydrolysis and ligand binding by a specific effect on the GTP-regulatory proteins (G-proteins) that regulate hydrolysis. This effect on hydrolysis was also observed in the RBL-2H3 cell (a cell that possesses the same type G-protein). Furthermore, the barbiturates inhibited the release of vasoactive substances from these cells at the same concentrations that were seen to inhibit hydrolysis. These data indicate that barbiturates may not only inhibit hydrolysis, but may also inhibit a functional cell response to hydrolysis, and may have an ubiquitous effect of any cell possessing the same type G-protein(s).

"Volatile Anesthetics and Hepatocyte Ca++ Homeostasis"
RO1 GM 38033 (Van Dyke, R.), Mayo Clinic

Historically, the effects of the volatile anesthetics on cell membranes have been considered to be nonspecific. While the volatile anesthetics may still produce anesthesia by nonspecific, general effects on membranes, there are also effects on specific sites within the membrane which are reflected in the pharmacodynamics of the volatile anesthetics. Dr. Van Dyke and coworkers have focused on the specific effects of the volatile anesthetics by studying the adrenoreceptor signal transducing system in rat hepatocytes. Their recent studies have led them to conclude that there is no general membrane effect on this system by the anesthetics but a site-specific effect.

Studies of the effects of the anesthetics on the kinetics of phenylephrine and prazosin displacement of C-14-prazosin revealed a very interesting pattern. It was found that halothane, enflurane, and isoflurane showed a dose-dependent decrease in B_{max} and an increase in K_b. The dose-dependent effects are in the clinical range of the anesthetics. These data suggest these anesthetics produce a decrease in affinity of the α_1-receptor for agonist and a decrease in receptor numbers. Similar results were obtained for the α_2- and β_2-adrenoreceptors, although these receptors are in much lower concentration on the hepatocyte plasma membrane than is the α_1-adrenoreceptor.

Using isolated rat hepatocytes in suspensions or primary cultures, the effects of halothane, enflurane, and isoflurane on intracellular messengers controlled by the adrenoreceptor was carefully examined in Dr. Van Dyke's laboratory. The release of intracellular calcium, cyclic AMP, and inositol phosphates have been measured. At clinical concentrations of anesthetics, there is a stimulation of calcium release. This was measured by three techniques: 1) calcium-45 release from calcium-45-loaded intact or saponin-treated hepatocytes; 2) aequorin-loaded hepatocytes; and 3) fura-2-loaded cells. All three techniques revealed a similar pattern of calcium release in response to the presence of the three anesthetics. Phospholipase C activity was found to be stimulated by clinical levels of the anesthetics and was temporally consistent with the release of calcium. In both cases, the stimulation occurred within seconds of the additions of the anesthetics. The effects of these anesthetics on cyclic AMP was found to be concentration-dependent, i.e. at low concentrations (less than 1 MAC where MAC is the minimum alveolar concentration of an anesthetic necessary to inhibit response to a noxious stimulis in 50 percent of subjects) the cyclic AMP synthesis was inhibited while at concentrations greater than 1 MAC the synthesis was stimulated.

In addition to the results summarized above, Dr. Van Dyke and his coworkers have made two observations of potential importance. First, a number of pitfalls exist in the use of fura-2 for studies involving hepatocytes and halothane. There have been reports that hepatocytes metabolize fura-2 rapidly and, while some metabolism was found, of greater consequence is the loss of fura-2 due to photobleaching. While this was known to occur, it was interesting to discover that halothane exacerbates the photobleaching presumably due to the formation of halothane free radicals. Neither enflurane nor isoflurane promoted photobleaching of fura-2 and neither are metabolized by way of free radicals. A second major observation relates to the resting levels of calcium in hepatocytes. In cells from animals pretreated with

phenobarbital to induce cytochrome P-450, the resting levels of calcium were found to be several-fold (3-4) greater than in cells from untreated animals. Using cytochrome P-450 levels as an indicator of induction, they have found a linear relationship between the amount of P-450 and resting levels of calcium.

BIORELATED CHEMISTRY RESEARCH OVERVIEW

Research in basic chemistry remains a fundamental area of interest for the Pharmacological Sciences Program. Basic studies in chemistry serve to uncover the patterns and rules by which biomolecules undergo covalent reactions or associate with a functional objective. This can be accomplished through direct studies of the biomolecules themselves, and this is becoming increasingly possible as advances in spectroscopy and other analytical techniques and in molecular biology open new avenues to their study. However, much of our information is still gained through model studies, using related, but more accessible systems. Indeed, organic synthesis in many instances can be viewed as a model study for reactions that occur in the body. Model host-guest chemistry has much to say about enzyme mechanisms and drug-receptor interactions.

An equally important benefit of research in basic chemistry is the many practical applications. As we learn more about the forces and mechanisms that operate in body chemistry, we also learn more about how to control or modulate those processes. Deciphering chemical mechanisms contributes valuable information for drug design. Advances in synthetic chemistry allow potential therapeutic drugs to be made. New biochemical probes and diagnostic agents are also derived from studies in basic chemistry. New synthetic or semi-synthetic reagents for carrying out delicate chemical transformations result from these studies. New biomaterials for implantable devices are discovered. New modes of therapy are discovered as drug design becomes more sophisticated. For example, we can now envision genes as a target of drug action. We are even developing new computer programs and systems to enhance and speed up this process.

The biorelated chemistry grants can be considered in four main groupings, one is natural products chemistry including isolation and structure elucidation, synthesis, and biosynthesis; the second is the study of biomedically significant chemical processes including organic and inorganic models of biological systems; the third is the study of the relationship between chemical structure and biological activity; and the fourth is synthetic and analytical methodology. No attempt has been made in this report to group the research highlights inasmuch as the divisions are blurred and many projects readily fall in several categories. However, the categories do serve to demonstrate the breadth of activities in biorelated chemistry.

Many isolation/structure elucidation studies in natural products chemistry have been focusing on the chemical defense mechanisms of plants and animals and the chemicals involved in symbiotic relationships among small organisms, plants, and animals. The bioactive substances discovered often show biological effects relevant for use in human health care. For example, one

study of the chemical defenses of sea corals is yielding compounds with anti-inflammatory properties. As soon as interesting new structures are discovered, synthetic chemists plot their synthesis. Since the previous biennial report, synthetic progress has been made on a number of biologically interesting natural products, such as the ginkolides which display significant effects on platelet activating factor and the pyrrolizidine alkaloids that show promising antiviral activity. Also as new structures appear, biosynthetic chemists are challenged to provide insight into the methods of their formation in the organisms. A clear trend in this area is the increasing use of molecular biology techniques to study biosynthesis. An additional benefit of this approach is the possibility of enhancing production of a useful chemical agent or directing the organism, through mutagenesis, to produce variants of the desired compound, in effect speeding up evolution to provide new bioactive structures.

The study of the structure and reactivity of metal-containing enzymes, ranging from P-450 enzymes to nitrogenase, is an area of substantial activity. Here, an array of advanced spectroscopic, X-ray, and magnetic techniques are being brought to bear on both the enzymes themselves and on model systems. Insights into enzyme mechanisms are being gained and new metallochemistry is being discovered. Noteworthy was the first X-ray diffraction characterization of a dioxygen complex of copper, and the study of sulfur/selenium compounds in intact erythrocytes. Singlet oxygen continues to command a great deal of interest, particularly in regard to its intermediacy in photodynamic therapy. The biorelated chemical processes most often being modeled are enzyme actions. Thus, molecular recognition studies, for example, host-guest chemistry, remains a growth area. The frontiers for abzymes, semi-synthetic enzyme mimics from antibodies, continue to be pushed back in terms of the range of reactions catalyzed and new approaches to catalysis being developed. Work on totally synthetic enzymic mimics is advancing, and the first fully designed, bimolecular enzyme mimic has been reported. Interestingly, work on enzyme mechanisms and models is revealing that some of the chemistry thought to be enzyme-catalyzed is operative without the enzyme, i.e., shikimate and vitamin B_{12} chemistry.

The trend in structure activity studies to involve computer graphics and energy minimization programs is continuing to expand. This area is marked by drug design efforts to improve agent design for enzyme inhibitors and receptor agonists and antagonists. An active area for some time has been the preparation of analogs of bioactive peptides. A new and strongly emerging theme is the design and synthesis of peptide mimics along the line of the morphine/enkephalin prototype, but with increasing sophistication, involving mimics of larger peptides with defined secondary structure. Peptide mimics, of course, hold promise of avoiding the uptake and metabolic liabilities of the natural peptides. Exploring the chemistry involved in covalent-reactive drugs, such as bleomycin and bicyclomycin, continue to offer insights for additional drug design. New prodrugs and other chemical drug delivery systems are being investigated. Another active area is agents that have DNA or RNA as a target.

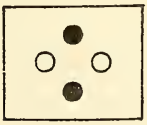
The development of synthetic methodology is also receiving the benefit of computers, aiding in synthesis design and the prediction of reactivity and stereochemistry. Biomolecules such as oligonucleotides, carbohydrates, and peptides are important targets of synthetic methodology development. Advances

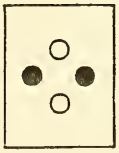
in the development of metal-based synthetic reagents is a strong area of interest. The synthesis and chemistry of medium-to-large rings also takes up a good deal of chemical effort. Perhaps the most significant trend in methodology development is the increased emphasis on enantioselective syntheses. This, of course, is of extraordinary significance to pharmacology and therapeutics, given the normally widely disparate biological activities of enantiomers. The term "chemzyme" has been coined to describe metal-based reagents capable of catalyzing organic transformations in a stereoselective manner at enzyme-catalyzed rates. Moreover, research findings in chemical methodology ultimately feed into all other areas of biorelated chemistry.

As researchers in pharmacology and anesthesiology probe deeper in their studies of the biological response to drugs, researchers in biorelated chemistry augment that effort by providing the chemical tools and chemical understanding necessary for biological characterization at the molecular level. The complementary efforts of investigators in the pharmacological sciences hold great promise for yielding basic discoveries that will translate into improved clinical care.

"Host Molecules that Complex and Catalyze"
R01 GM 12640 (Cram, D.), University of California, Los Angeles

Dr. Cram's research group has designed, synthesized, and studied a new type of compound which strongly self-associates in organic solvents without any hydrogen bonds, ion pairs, or metal ligation driving forces. Binding free energies well in excess of -7 kcal mol^{-1} have been demonstrated for dimer formation in solvents such as acetone or ethyl acetate. A crystal structure of one dimer shows the presence of over 56 atom-to-atom contacts across the surface common to each complexing partner. These interatomic distances are in the 3.3 to 3.8 Å range needed for operation of dipole-dipole attractive forces. Each molecule of the compound class contains a rigidly organized, nearly flat face of about 15 Å by 20 Å dimension, composed of assemblies of aromatic rings. Two methyl groups protrude from the surface, like doorknobs at 12 and 6 o'clock, and two holes complementary to these knobs are found at 3 and 9 o'clock, as illustrated in schematic A. Schematic **B** is identical to A except that it has been rotated 90°. When A and **B** contact one another


A


B

face to face, the four knobs lock into the four holes to produce dimer AB, whose faces cannot move with respect to one another. The entropic cost of forming such a highly organized dimer is paid for by the release upon complexation of 10 to 20 partially organized solvent molecules that solvate the faces of the two monomer components. To make the monomers and dimer soluble, four long, flexible groups are attached to the side opposite the complexing face in each molecule.

This unique family of compounds is of great interest and importance for several reasons. The binding free energies holding these dimers together in solution can be as low as -2 kcal mol^{-1} in nonpolar solvents such as toluene. As solvents become more polar in the sequence toluene, chloroform, acetone, ethyl acetate, nitromethane, and methanol, the binding free energies increase monotonically to values well in excess of -7 kcal mol^{-1}. In water, the values should be enormous. Thus, large organic solvaphobic effects are subject to systematic scrutiny for the first time. Multiple functional groups that catalyze organic reactions cooperatively can be attached to the monomer units and positioned properly by such dimerizations. Two monomer molecules can be covalently attached to one another through their solubilizing appendages to form new molecules that contain two faces oriented 180° from one another. Such compounds should self-assemble to form a new kind of highly linear polymer of very unusual properties, potentially useful to materials science. This work is inspired by, and makes use of, the principles of molecular recognition in complexation, which is central to the biological control mechanisms of physiological chemistry. Self-assembly of polymers held together by noncovalent forces are ubiquitous to nature in systems as widely different as enzyme-fine structures, double and triple helices of the genes, and the compositions of lipophilic membranes.

"Treatment Planning in Photodynamic Tumor Therapy"
R01 GM 20117 (Grossweiner, L.), Illinois Institute of Technology

Photodynamic therapy (PDT) is an experimental treatment for malignant tumors based on the combined action of strong red light and a porphyrin drug, HPD, that localizes selectively in neoplastic tissue after intraveous injection. Since the therapeutic light must be absorbed by HPD, the light dose depends on the HPD concentration in the tumor, as well as the tumor dimensions, optical properties, and the method of light delivery.

PDT is presently an experimental therapy and is being carried out under rigid protocols that facilitate statistical evaluation of safety and efficacy. However when PDT is fully approved, treatment planning will be required for each case to optimize the probability of effecting a cure. Dr. Grossweiner is developing methods for analytical PDT treatment planning and comparing the predictions with on-going clinical trials and laboratory studies on model systems.

The clinical objective in PDT is to deliver a light dose which ensures that all tumor regions absorb the threshold energy density required for tumor eradication. Energy absorption in tissues cannot be calculated by usual spectrophotometric methods because multiple light scattering leads to highly turbid, almost opaque materials. The methods of tissue optics can be employed for this purpose, based on approximate radiative transfer theories. The spatial distribution of radiant power density in an idealized tumor model is calculated for various modes of light delivery including external and interstitial optical fibers. A treatment plan is developed by determining the light dose required to achieve the threshold absorbed energy density at the tumor boundaries. The input parameters are estimated from literature data and the results of clinical trials. The current form of the model is being

evaluated by retrospective comparisons to clinical trials on PDT of head and neck squamous cell carcinoma being carried out at Ravenswood Hospital Medical Center in Chicago.

Another approach to PDT dosimetry being employed by Dr. Grossweiner involves studies on tissue phantoms that mimic the initial photochemical stages of PDT. A typical phantom is comprised of defined light-scattering entities, a photosensitizer, and a biological target. Extensive studies have been made on aqueous suspensions of polystyrene microspheres in the presence of HPD and a colorless enzyme. The experimental rate of photosensitized enzyme inactivation is compared with calculations based on tissue optics theory. Other tissue phantoms utilize liposomes and resealed red blood cell membranes, in which case membrane damage provides a convenient damage assay. The detailed calculations require values of the relevant optical constants. A computerized diffuse optics spectrophotometer has been constructed for this purpose that measures the diffuse reflectance and transmission of layer samples at 1 nm intervals in the near-ultraviolet, visible, and near-infrared regions. These investigations provide an independent test of tissue optics theory and facilitate the evaluation of wavelength effects, light penetration, HPD photobleaching and other photophysics-chemical factors common to these tissue phantoms and PDT.

"New Bioorganic Methods for Studying Membrane Structure"
RO1 GM 21457 (Menger, F.), Emory University

Phospholipids spontaneously form molecular bilayers when mixed with water. The layers, only two molecules thick, can be prepared in the form of spherical shells or liposomes that surround and encase small volumes of water or solutions containing drugs. Hence, liposomes are being studied intensively for their ability to serve as drug delivery systems.

Dr. Menger's laboratory has synthesized over two dozen new phospholipids containing branches (i.e. methyl, butyl, phenyl, etc.) at various locations along the phospholipid chains. Owing to these branches, the "unnatural" lipids are unable to pack as efficently into the liposomal bilayers. As a result, the liposomes leak their contents at rates that depend on the size, location, number, and polarity of the substituent branches on the chains. The possibility now exists of controlling via organic synthesis the rate at which drugs can be delivered to a patient, thereby averting the concentration bursts that arise from taking periodic large doses.

The new phospholipid compounds have been characterized by a variety of modern experimental methods in an effort to understand their behavior when incorporated into liposomes. One such method used by the Menger research group is differential scanning calorimetry. When lipid suspensions are slowly heated, the chains eventually become disordered at a sharp temperature called the transition temperature. Transition temperatures and associated thermotropic properties provide information about the packing properties of the lipid assemblage. For example, it was found that methyl substitution at carbons 4, 10, and 16 of the second lipid chain gives transition enthalpies of 8.8, 3.7, and 9.8 kcal/mol, respectively. This remarkable effect of methyl groups in the middle of the chains (carbon 10) was reconciled with a membrane model in which chain mobility is restricted in the first half of the chain

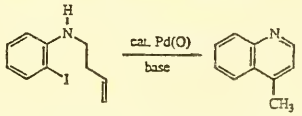
after which motional freedom and disorder increase uniformly up to the terminal methyl. Additional experiments on monomolecular films, using a Langmuir surface balance, confirm this picture. On the basis of this work, it was predicted that a methyl at the center of the chain should impart a maximal effect on transport through liposomal walls, and such was found to be the case. Thus, diffusion from one side of the bilayer to the other side occurs most efficiently when the lipid packing is disrupted by substituents at the middle portion of the chains.

The library of synthetic lipids was tested for biological activity. In work carried out collaboratively with Professor J. F. Kuo, it was discovered that several branched-chain phospholipids can, depending upon the size of the lipid substituent, either stimulate or impede the activity of the important enzyme, protein kinase C. For example, the 4-butyl and 8-butyl derivatives increase activities to a level comparable to that seen with diolein, a diacylglycerol derivative often cited as a classic protein kinase C activator. A second enzyme, phospholipase-A2, is affected by the new phospholipids. Normally, this enzyme cleaves the second fatty acid off phospholipids. But when alkyl groups, particularly large ones like n-butyl, are located in the first half of the chains, the phospholipid becomes resistant to cleavage by the enzyme. More importantly, the synthetic phospholipid binds to the enzyme and prevents it from cleaving a normal substrate. In other words, chain-substituted lipids serve as inhibitors for phospholipase-A2. Developing such inhibitors is a major goal of current chemo/pharmacological research because phospholipase-A2 catalyzes the release of arachidonic acid, a precursor for prostaglandins that are involved in inflammatory processes such as arthritis.

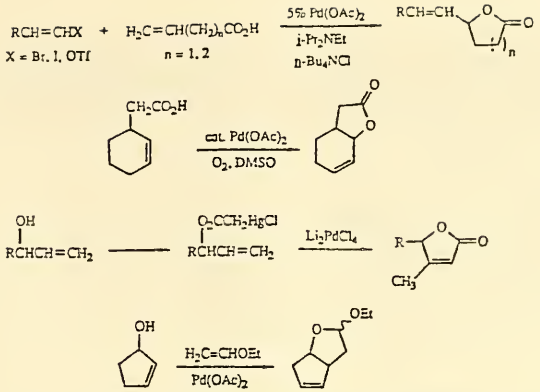
"Palladium Approaches to Heterocycles and Prostaglandins"
RO1 GM 24254 (Larock, R.), Iowa State University

Dr. Larock's research efforts have been devoted to the development of new synthetic routes to heterocycles and prostaglandins which possess potential biological and pharmacological activity. His research group is taking advantage of the ease with which palladium catalyzes new carbon-carbon bond formation while accommodating a wide range of important organic functional groups.

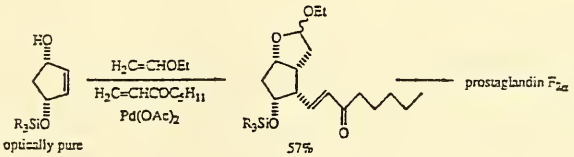
Recent work has provided new routes to indoles, indolines, quinolines, isoquinolines, isoquinolones, and benzofurans.



New routes to lactones and acetals have also been developed.

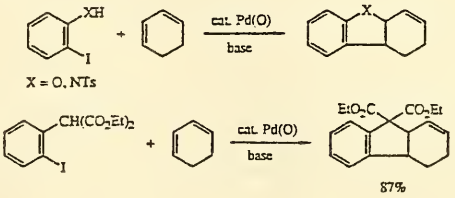


The latter reaction provides a novel, highly efficient route to the human hormones known as prostaglandins. Work on a variety of prostaglandin

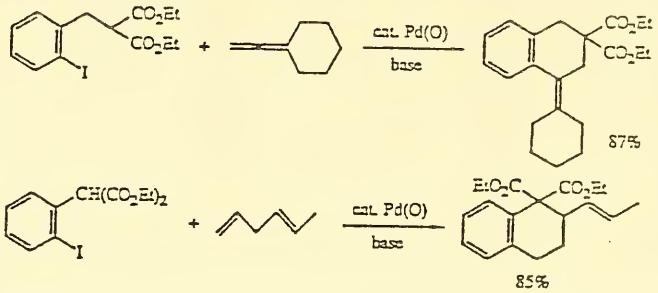


syntheses has also been completed. For example, vinylpalladium chemistry has provided a remarkably efficient new approach to prostaglandin E_2.

In recent years the Larock group has also become very interested in new annulation chemistry. New methods for both heteroannulation and carboannulation of allenes, conjugated and non-conjugated dienes, and alkynes have recently been discovered.



The work recently developed by Dr. Larock and his coworkers greatly simplifies the synthesis of heterocycles, carbocycles and prostaglandins. It should also prove useful in the synthesis of a wide variety of other pharmacologically interesting compounds.



"Studies in the Shikimate-Chorismate Pathway"
R01 GM 28965 (Bartlett, P.), University of California, Berkeley

The shikimate-chorismate pathway, which plays a key role in the biosynthesis of aromatic compounds, involves a number of enzyme-catalyzed transformations with novel mechanisms. Recent work in Dr. Bartlett's group has served to delineate important aspects of these mechanisms and the role that the respective enzymes play in inducing their reactions. These studies have led to the design and synthesis of potent inhibitors for a number of the enzymes, and even provided the basis for the genesis of non-enzymic, catalytic proteins.

The reaction mediated by dehydroquinate (DHQ) synthase, one of the first enzymes in the pathway, involves a complex sequence of steps culminating in the formation of a carbocylic ring via an intramolecular aldol condensation. Formation of the immediate precursor to this step via a non-enzymic reaction allowed the Bartlett group to demonstrate that this cyclization does not require enzymatic catalysis. This result suggests that the role played by the enzyme is considerably less complex than assumed previously, and focuses attention on the inherent chemical behavior of the substrate itself.

A step in this pathway which has key economic importance is that catalyzed by the enzyme 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase. This enzyme mediates a vinyl ether exchange reaction and is the target of the herbicide Glyphosate (sold under the tradename "Round Up"). By devising stable mimics of the unstable intermediate in this transformation, Dr. Bartlett and his coworkers have discovered an exceedingly potent series of inhibitors for this enzyme, with inhibition constants in the nanomolar range. Their ability to assign the three-dimensional structure of these compounds has also allowed them to propose the stereostructure of the transiently formed, unstable

intermediate in the enzymatic process. The synthesis of fluorinated analogs of EPSP, which is the product of EPSP synthase, has enabled the mechanism of the reverse reaction to be probed, as well as the mechanism of the subsequent reaction in the sequence, which is catalyzed by chorismate synthase.

Chorismate synthase induces the elimination of an allylic phosphate to generate the cyclohexadiene moiety of chorismate. The mystifying element to this mechanism is the enzyme's requirement for an oxidation-reduction cofactor, in spite of the fact that neither oxidation nor reduction is necessary according to the 'normal' mechanism of the reaction. To probe this question, the Bartlett group has synthesized isotopically labeled derivatives of the substrate, to see if the elimination reaction is a concerted or non-concerted process. These studies have already suggested that the transformation proceeds via radical intermediates, which would be unprecedented for the mechanism of an elimination reaction.

The design and development of inhibitors of the enzyme chorismate mutase has been a central part of this research project, and indeed a focus of a number of research groups over the years. This enzyme has been an intriguing target for inhibition because it is one of the few whose reaction has an observable, non-catalyzed counterpart, and because this transformation is perhaps the only example of a Claisen-type rearrangement in primary metabolism. Dr. Bartlett's laboratory has reported the design, synthesis, and enzymatic evaluation of a number of molecules designed to mimic the transition state conformation of the substrate chorismate, including that of the most potent inhibitor known to date. The behavior of this inhibitor, which is an oxabicyclic diacid, has helped to refine ideas about the balance between conformational and electronic stabilization of the transition state in the enzymatic reaction. Perhaps most interestingly, the Bartlett inhibitor has been used as a hapten for the development of monoclonal antibodies with catalytic activity, that is, it has served as the starting point in a study to develop antibodies which are capable of catalyzing the rearrangement.

"Structure of Protein-Bound Bile Pigments"
R01 GM 28994 (Rapoport, H.), University of California, Berkeley

The phycobiliproteins serve as light-harvesting proteins in cyanobacteria and red algae. The visible light-absorbing groups on these proteins are open - chain tetrapyrroles derived from biliverdin which are covalently attached to the polypeptide chains. The spectroscopic properties of particular phycobiliproteins are dependent on the chemical nature of the tetrapyrroles (bilins) attached to it. Drs. Rapoport and Glazer and their associates developed procedures for the isolation of the tetrapyrrole prosthetic groups in their peptide-linked form. By studying such bilin peptides by high resolution NMR techniques they were able to establish both the protein-bilin linkage(s) and determine the structure of the bilin. These studies revealed that the bilins were all attached to the proteins through either one or two thioether linkages to cysteinyl residues. Four chemically distinct bilins are attached to the cyanobacterial and red algal phycobiliproteins. Determination of the structures of these peptide-linked bilins led to the finding that they are isomers differing only in the arrangement of double bonds. The absorption spectra of the family of pigments generated in this simple manner span the total range of visible wavelengths.

Recently, this group has completed the total synthesis of phycocyanobilin covalently attached to cysteine. This has been achieved with control and assignment of the absolute stereochemistry at all asymmetric centers. Projection of this synthetic methodology should make available synthetic phycocyanobilipeptides that mimic the properties of the natural materials.

The ready oxidizability of open-chain tetrapyrroles observed in the course of studies of the biosynthesis of the phycobiliproteins led to the suggestion that bilirubin, an end product of heme degradation in man, may be a physiologically important antioxidant. Collaborative studies provided strong evidence of this role for bilirubin. Bilirubin is transported in the blood bound to serum albumin. In particular, it was demonstrated that unsaturated fatty acids bound to albumin were spared from oxidation by peroxyl radicals at the expense of the oxidation of bilirubin to biliverdin. Since biliverdin is converted to bilirubin by biliverdin reductase in vivo, any such oxidized bilirubin would be regenerated.

In another outgrowth of the research on the phycobiliproteins, attack by reactive oxygen species has been shown to lead to a decay in the intense fluorescence of the phycobiliprotein phycoerythrin. This phenomenon has been exploited to develop a versatile new screen for small molecules and macromolecules that may function as biologically relevant protective agents against free radical damage.

"Analogs of S-Adenosylmethionine"
RO1 GM 29332 (Borchardt, R.), University of Kansas

In the past several years, Dr. Borchardt and his colleagues have shown that S-adenosylhomocysteine (AdoHcy) hydrolase (EC 3.3.1.1) is an attractive target for the design of antiviral agents. This cellular enzyme catalyzes the reversible hydrolysis of AdoHcy to adenosine and homocysteine, which are normally removed rapidly by further cellular metabolism. However, inhibition of cellular AdoHcy hydrolase results in intracellular accumulation of AdoHcy, a significant increase in the intracellular AdoHcy/S-adenosylmethionine (AdoMet) ratio and subsequent inhibition of AdoMet-dependent methylation reactions such as those essential for viral mRNA maturation.

Most recently, Dr. Borchardt's efforts to design more potent and more specific inhibitors of AdoHcy hydrolase have focused on developing a computer graphics model of the active site of this enzyme. Since an X-ray crystal structure for AdoHcy hydrolase does not exist, Dr. Borchardt has used the crystal structure of lactate dehydrogenase (LDH). LDH was chosen because its primary sequence shows considerable homology to the primary sequence of AdoHcy hydrolase. The active site model of AdoHcy hydrolase, including the NAD^+ binding domain, was obtained by systematically mutating the structure of LDH. The nonidentical amino acid residues in LDH were replaced using MUTAR, which was developed by Dr. Borchardt's collaborator Dr. A. Vedani. The structure was refined by using AMBER, which works in a cartesian coordinate concept and relaxes all atoms of the protein, and YETI, which works in an internal coordinate concept and does not alter the backbone configuration. This potential model of the active site of AdoHcy hydrolase has been used to explain the relative activity

of known inhibitors and as a tool to design new and more potent inhibitors of this enzyme.

Dr. Borchardt's laboratory has also been interested in elucidating the mechanism by which carbocyclic nucleosides (e.g., neplanocin A) inactivate AdoHcy hydrolase. Results of their research have shown that neplanocin A inactivates the enzyme by a "cofactor-depletion" mechanism, involving conversion of the enzyme from the NAD^+ form (catalytically active) to the NADH form (catalytically inactive). The 3'-keto nucleoside arising from this inactivation process binds tightly to the NADH form of the enzyme, but it does not become covalently bound to the protein.

In collaboration with Dr. Erik De Clercq's laboratory, Dr. Borchardt has also shown that synthetic analogs of neplanocin A are broad spectrum antiviral agents. The antiviral activity of these AdoHcy hydrolase inhibitors can be increased substantially by coadministration of homocysteine. The mechanism responsible for this synergism was shown to involve an inhibitor-insensitive form of AdoHcy hydrolase and its ability to catalyze the biosynthesis of AdoHcy from endogenous adenosine and exogenously administered homocysteine. The function of this inhibitor-insensitive form of AdoHcy hydrolase is currently under investigation.

"Chemistry of Drug Action of Serine MAO Inhibitors"
RO1 GM 32634 (Silverman, R.), Northwestern University

Monoamine oxidase (MAO) is one of the enzymes responsible for the degradation of various neurotransmitters. Compounds that inactivate MAO exhibit antidepressant activity new classes of MAO inactivators are highly desirable for testing as antidepressant drugs. Dr. Silverman has been working to understand the mechanism of MAO and to use this knowledge to design MAO inhibitors.

Over the years Dr. Silverman has provided strong indirect evidence for a reaction involving radical (one-electron) intermediates, but now he has observed the first electon spin resonance spectrum (an experiment that detects one electron intermediates) for this reaction. This was made possible through the design of a molecule that could act as a substrate for the enzyme but for which the intermediate species generated was a much more stable radical (although still a short-lived species). In the presence of another compound that can react with short-lived radicals and produce long-lived radical species, it was possible for Dr. Silverman to make this first observation of a radical produced by MAO.

He is also using mechanistic studies to improve drug selectivity. MAO exists in two different enzyme forms known as MAO A and MAO B. The earlier MAO inactivators reacted with both forms of the enzyme, producing not only the desired antidepressant effect but also undesirable cardiovascular side effects unless the diet was restricted to eliminate tyramine-containing foods. The sequences of amino acids that constitute each of these enzyme forms are known, but the composition of the substrate binding site is not known. Using purified MAO A and MAO B Dr. Silverman has discovered a potent inactivator of both forms of the enzyme that attaches to one amino acid residue (a cysteine) in each form of the enzyme. This inactivator has been synthesized with a

radioactive tag so that he could determine that only one molecule of inactivator became bound to the substate binding site. He is now ready to degrade the enzymes containing the radioactive tag and identify the structure of the peptide at the substrate binding site. This should provide an important basis for the design of highly selective inactivators of each form of the enzyme.

In addition, Dr. Silverman has designed two new classes of MAO inactivators, (aminoalkyl)trimethyisilanes and (aminoalkyl)trimethylgermanes and has synthesized radioactive analogs of two of these compounds. Both compounds exhibit unexpected, but intriguing, chemistry with MAO which should make it possible to identify the structure of part of the substrate binding site.

"Large-Ring and Helical Complexing Agents for Metal Ions"
R01 GM 32937 (Bell, T.), State University of New York at Stony Brook

The relationship between ionophore structure and useful properties, such as strong binding, rapid exchange, and selectivity, is poorly understood. In most synthetic analogs of ionophores, such as crown ethers, cryptands, and spherands, molecular rigidity, and metal encapsulation are linked, leading to enhanced binding strength and selectivity. Dr. Bell's research group has synthesized "torands," which are extremely rigid yet nonencapsulating complexing agents for metal ions. A key example of this new ligand family has been found to form unusually stable complexes with alkali metal ions such as sodium and potassium.

Torand complexes exchange rapidly despite their high stability. This unique combination of properties is attributed to the openness of the metal-binding cavity, which is defined by several rigidly converging pyridine rings. Potential applications of complexing agents with similar properties include sequestration or regulation of metals in biological fluids and preparative separation of biomedically important metals, including their isotopes.

A second line of research in Dr. Bell's laboratory explores the application of torand "hexagonal lattice" molecular architecture to the design and synthesis of artificial receptors for organic molecules. One example is a complexing agent that extracts urea from water. This approach may eventually lead to a new method for continuous optical measurement of BUN levels in blood serum. Other new receptors bind guanidine and benzamidines and have potential for application to in vivo determination of antiprotozoal drugs such as pentamidine.

"Mitochondrial Iron-Sulfur Proteins"
R01 GM 33806 (Johnston, M.), University of Georgia

Iron-sulfur clusters greatly outnumber cytochromes in the mammalian mitochondrial respiratory chain. However, their spatial organization as well as the properties and functions of individual clusters are poorly understood in comparison with cytochromes. Dr. Johnson's group is using a variety of spectroscopic techniques at cryogenic temperatures, notably resonance Raman, magnetically induced circular dichroism, and electron paramagnetic resonance,

to investigate the structural, electronic, and magnetic properties of the iron-sulfur clusters in enzymes and proteins isolated from beef heart mitochondria. These include NADH dehydrogenase, succinate dehydrogenase, Rieske protein, aconitase, and electron transfer flavoprotein dehydrogenase.

Previous reports from Dr. Johnson and coworkers have established the presence of three distinct types of redox-active iron-sulfur clusters in mammalian succinate dehydrogenase: one each of $[2Fe\text{-}2S]^{1+,2+}$, $[3Fe\text{-}4S]^{0,1+}$, and $[4Fe\text{-}4S]^{1+,2+}$ clusters. During the past year, research efforts on this enzyme have focused on elucidating the spatial arrangement and subunit location of these clusters and the mechanism of electron flow through the intermediate, Complex II. One particularly fruitful approach has involved parallel activity assays and spectroscopic studies of genetically modified forms of fumarate reductase from Escherichia coli. This enzyme, which catalyses the same reaction only in the reverse direction, has been shown to have the same subunit and iron-sulfur cluster composition as mammalian succinate dehydrogenase. The subunit location of the constituent iron-sulfur clusters has been established by electron paramagnetic resonance studies of whole cells and whole cell extracts of a fumarate reductase deletion mutant with plasmid amplified expression of discrete subunits, parts of subunits, or groups of subunits. Information concerning the specific residues involved in ligating the clusters and their role in the catalytic mechanism has come from studies of isolated samples of mutated enzymes in which specific cysteine residues have been replaced by serines. Spectroscopic studies of these site-specific mutants have revealed structural modification or selective deletions of individual clusters.

The results have shown that the [2Fe-2S] and [3Fe-4S] clusters reside exclusively in the iron-sulfur protein subunit and have identified the coordinating cysteine residues. Of these two clusters, only the [2Fe-2S] cluster appears to be essential for enzymatic activity in vitro. Mutants in which this cluster is absent or has modified spectroscopic properties show a complete or partial loss of fumarate reductase activity. In contrast, mutants deficient only the [3Fe-4S] cluster show no significant decrease in fumarate reductase activity in vitro using artificial electron donors, but are no longer capable of accepting electrons from the physiological donor in vivo. Hence, while similar clusters in numerous other enzymes and proteins have been shown to be isolation artifacts, the [3Fe-4S] clusters in succinate dehydrogenase and fumarate reductase appear to have a physiological role in mediating electron transfer to or from the quinone pool. The results obtained thus far are less definitive with respect to the location and function of the low potential [4Fe-4S] cluster, but are most readily interpreted in terms of this center residing exclusively in the flavoprotein subunit or bridging between the flavoprotein and iron-sulfur protein subunits.

"Non-ionic Liposomes and Redox Activated Cation Pumps"
R01 GM 33940 (Echegoyen, L.), University of Miami

During the past 2 years, Dr. Echegoyen and his colleagues have discovered and developed a novel type of vesicular structure based on newly synthesized compounds. The compounds contain a lipophilic (or 'greasy') tail, usually a cholesterol or a derivative, and a cyclic polyether (crown ether) as the polar head group. Upon dispersion in water the molecules aggregate in pretty much

the same way as natural phospholipids do to form cell membranes. The resulting structures are called niosomes, short for non-ionic liposomes. Since crown ether groups are found on the outside and inside surfaces of these vesicles, and since these groups are capable of strongly binding metal cations such as sodium and potassium, these niosomes have potential applications in chemical catalysis, cation transport, and drug delivery. Many of these interesting properties are currently under investigation.

On a related, but independent development, lipophilic crown ethers (which contain anthraquinone substituents) have been synthesized. These molecules can be electrochemically or chemically reduced to their corresponding anions, which are stable at neutral pH, in order to enhance their cation binding ability. The acquired negative charge induces the formation of intramolecular ion pairs, thus increasing the cation binding ability of the ligand by several-fold. Dissolved in lipophilic media, these molecules are able to efficiently extract metal cations from a water phase in contact with the lipophilic one, especially when reduced at the interface. The resulting lipophilic cation complexes can then be easily oxidized at another interface to release the cations. The net result is the pumping of cations from a donor to a receiving phase, using a redox gradient across the membrane phase. The redox gradient has been generated electrochemically, and efforts to do it via chemical means are currently under way.

"Studies of Synthetic Metalloenzyme Mimics"
RO1 GM 34841 (Burrows, C.), State University of New York at Stony Brook

Molecular recognition chemistry in conjunction with the chemistry of transition metal catalysts has led to the development of new synthetic molecules capable of biomimetic action. The long-term objectives of this research are the identification of new catalysts for selective transformations of organic substrates and an understanding of the fundamental requirements for catalysis within an enzyme or a synthetic molecular recognition system. These results will be applied to the preparation of new biologically active materials or the degradation of naturally occurring biopolymers.

Dr. Burrows' group has discovered a series of macrocyclic nickel(II) complexes capable of catalyzing hydrocarbon oxidation with iodosylbenzene or hypochlorite as terminal oxidant. The ligands used include polyamine (cyclam) and polyamide-type macrocycles. Studies of the mechanism of oxidation indicate numerous similarities to cytochrome P-450 model systems. In particular, reaction conditions similar to those used with iron and manganese porphyrins leads, in the present case, to turnover rates of 200-300 per minute for alkene epoxidation--similar to the rates observed for a P-450 enzyme itself. The incorporation of a nickel-binding site into a bifunctional molecular receptor has also been achieved. This catalyst contains both the nickel(II) catalytic site and two crown ether moieties for complexation of a diamine substrate. Substrate oxidation studies are in progress in order to determine the intramolecular advantage in catalysis as well as the regio- and stereoselectivity. In the area of biopolymers, Dr. Burrows' group is investigating DNA damage by cationic Ni(II)-polyamine complexes using similar ligands and methodology.

In a second area, Professor Burrows' group has synthesized new dinuclear Ni(II) and Cu(II) complexes in order to probe possible mechanisms for amide or urea hydrolysis. The enzyme urease makes use of two active site Ni(II) ions in order to enhance the rate of urea hydrolysis by 14 orders of magnitude. As a mimic of this enzyme, macrocyclic Schiff base ligands containing two metal binding sites at fixed distances of about 6 Å have been prepared, and crystal structures have been obtained. Preliminary studies of ester hydrolysis show moderate rate enhancements.

"Stereocontrolled Amine and Peptide Analog Synthesis"
RO1 GM 35466 (Hopkins, P.), University of Washington

The eggs of a variety of marine invertebrates contain high concentrations of a family of amino acids called ovothiols. Present at concentrations of several millimolar, these substances may well be the most abundant small molecules in these eggs. Recent work in the laboratories of Dr. Hopkins and collaborator Dr. B. M. Shapiro has resulted in determination of the structures of these molecules, provided a source through chemical synthesis, yielded insights into the unique oxidation-reduction reactions of ovothiols, and has led to the proposal of possible roles of these substances in the egg and developing embryo.

The groups of Dr. Hopkins and Dr. Shapiro have reported that the ovothiols are structurally related to the amino acid histidine, and contain additionally one or more methyl groups and a single thiol substituent. A closely related family of substances of marine origin had previously been discovered and their structures studied by a team in Italy. Comparison of a sample isolated by the Italian group from the Bay of Naples to a sample isolated by the American team in the Strait of Juan de Fuca revealed the substances to be identical, and resolved a minor structural error on the part of the former group. In the ovothiols, the thiol functional group is attached directly to the aromatic, imidazole ring, making these molecules structurally unique among natural products.

To both prove the structures proposed for ovothiols and to provide a source for future studies, the preparation of these compounds by chemical synthesis was studied. A multistep pathway which provides optically active ovothiols A and C or multigram quantities of ovothiols A, B, or C was developed. These syntheses required the development of a new method of imidazole synthesis which is applicable to highly substituted imidazoles.

Although imidazole and thiol functional groups are not unique among naturally derived substances, their direct attachment to one another results in both structural and chemical novelty. The ovothiols are perhaps misnamed, because potentiometric titration of the thiol-substituted imidazole ring revealed that at physiological pH the predominant form is in fact a zwitterion possessing a thiolate anion and imidazolium cation.

The answer to why marine invertebrate eggs contain such extraordinary concentrations of ovothiols has not been proven, but a clue may be provided by the ease with which ovothiols are oxidized to the corresponding disulfides. In this property, they are significantly more reactive than glutathione.

Dr. Shapiro's group has measured the reactivity of ovothiols toward hydrogen peroxide and found it to be several times that of another abundant thiol, glutathione. They have proposed that ovothiols may play the role of scavenger of hydrogen peroxide in the developing embryo. An early biochemical event of marine embryos is a respiratory burst which includes the production of sizable quantities of this hydrogen peroxide for the specific purpose of oxidatively crosslinking the proteins present on the embryo surface. Ovothiols may protect the embryo from oxidative stress imposed by untoward reactions of hydrogen peroxide. Dr. Hopkins' group has noted that the ovothiols are quite active as scavengers of free radicals, presumably repairing these radicals and becoming themselves oxidized to the disulfide form. They have studied the scavenging of such radicals as Fremy's salt, galvinoxyl, and photogenerated tyrosyl radicals. In all cases, ovothiols were superior to glutathione. Because hydrogen peroxide in combination with various metal ions can produce highly reactive free radical species capable of oxidizing the cell's contents, this dual role of ovothiol as hydrogen peroxide and free radical scavenger may make it well suited to join vitamins A and C as nature's biological antioxidants.

"Asymmetric Synthesis and Its Applications"
RO1 GM 35879 (Masamune, S.), Massachusetts Institute of Technology

A current central issue of research in organic chemistry concerns the development of new strategies and methods for the enantio- and diastereoselective construction of new bonds through the process of asymmetric synthesis. Several nonenzymatic synthetic methods now exist to prepare, even on an industrial scale, numerous substances of physiological significance with near perfect enantioselection which was achievable earlier only through the use of enzymes. New asymmetric syntheses demand the design and preparation of homo-chiral reagents or catalysts. Dr. Masamune and coworkers have found that (*R*,*R*)- and (*S*,*S*)-2,5-dimethyl-1-boracyclopentanes (either one being referred to below as 1) exhibit remarkable enantioselection (>100:1) in hydroboration, ketone reduction, the aldol reaction, and allyl- and crotylboration.

Hydroboration. Reactions of symmetrically disubstituted *Z*- and *E*-alkenes and also trisubstituted alkenes with 1 followed by standard oxidative workup provide the corresponding alcohols with a minimum of 95 percent ee and with 99 percent ee or more in most cases. Reagent 1 constitutes the first logically designed hydroborating agent, and perhaps most importantly, it provides crucial information as to the transition-state geometry of the reaction, as its course is readily analyzed with the conformationally fixed chiral 1. A theoretical evaluation has also been made. Ketone reduction. In the presence of 0.2 equivalent of methanesulfonic acid, 1 reduces unsymmetrical dialkyl ketones with near perfect enantioselection. It even distinguishes between the two alkyl groups in the reduction of methyl ethyl ketone. The mechanistic course of this catalytic reaction has been elucidated and the proposed transition state is consistent with the result obtained from computational studies. Aldol Reaction. The enolate derived from *S*-3-(3-ethyl)pentyl propanethioate with (*S*,*S*)-2,5-dimethyl-1-boracyclopentane *B*-trifluoromethanesulfonate 2 and diisopropylethylamine undergoes aldol reactions with a variety of aldehydes to provide anti-2-methyl-3-hydroxy-carbonyl compounds with high diastereoselection (anti/syn >30:1) and enantioselection (>97 percent ee for the anti-isomer). The products represent

an important structural unit of polyketide-type natural products and many other compounds. The reaction is one of a few enantioselective processes that achieve the carbon-carbon bond formation of this type. Reagent **2** is readily prepared from **1** with trifluoromethanesulfonic acid. Allyl- and Crotylboration. These reactions are synthetically equivalent to aldol reactions. Reactions of *B*-allyl-1-boracyclopentane, derived from **1**, with representative aldehydes provide the corresponding homoallylic alcohols of 85-93 percent ee. In a similar manner B-(Z)- and (E)crotyl-(R,R)-2,5-dimethyl-1-boracyclopentanes, upon reaction with aldehydes, lead to the predominant formation of the syn- and anti-2-methyl-1-hydroxyl products, respectively, with an average selection of 20:1. The ee's of the major product range between 86-97 percent for the former and between 95-97 percent for the latter.

A detailed mechanistic analysis of the above four reactions involving **1** and its related reagents of C_2 symmetry has led to the investigation on the asymmetric induction effected by the 2-monosubstituted boracyclopentanes of C_1, symmetry which are more readily prepared than those of C_2 symmetry. Somewhat unexpectedly, it has been discovered that B-allyl-2-(trimethylsilyl)-1-boracyclopentane undergoes allylboration with aldehydes with a >96 percent ee in most cases, finally achieving satisfactory stereoselection for the acetate addition or its equivalent.

"Mechanisms of Biosynthetic Formation of Deoxy Sugars"
RO1 GM 35906 (Liu, H.), University of Minnesota

The deoxy sugars, long known as an important class of carbohydrates, occur widely in nature. Inspired by the intriguing biological activities possessed by the deoxy sugars and the perplexing steps involved in their biotransformation, Dr. Liu and coworkers have initiated a study to examine their biosynthesis at the molecular level. Efforts over the past few years have been concentrated on the mechanistic study of the formation of 3,6-dideoxyhexoses which are found only in the lipopolysaccharides of gram-negative bacteria. These unusual sugars have considerable biological significance, as they have been shown to contribute to the serological specificity of many immunologically active polysaccharides. Namely, they are the dominant antigenic determinants of the surface antigen of Gram-negative bacteria. Based on the pioneering efforts of Strominger and his coworkers, the pathway for the biosynthesis of ascarylose, a 3,6-dideoxy-L-arabino-hexose, had been put forward. The key reaction of this proposed sequence is the C-3 deoxygenation step catalyzed by enzymes E_1, a pyridoxamine-5'-phosphate (PMP) linked enzyme, and E_3, a NADPH dependent catalyst, both of which had been purified from Pasturella pseudotuberculosis. Although the catalytic roles of these enzymes have been well defined, the intimate mechanism of these steps is still disputable. In an effort to clarify these mechanistic ambiguities, this group has isolated an 'E_3 equivalent' from Yersinia pseudotuberculosis which is known to have ascarylose as the nonreducing terminal sugar in its lipopolysaccharide structure. Since the purified enzyme functions not only as a 3,4-glucoseen reductase but also as a NADH oxidase, this study was initiated by characterization of the catalytic properties of this enzyme as a NADH oxidase.

The purified enzyme contains no metals and consists of a single polypeptide chain with a molecular weight of 41,000. Its UV-VIS spectrum is that of a simple polypeptide with an absorption maximum around 280 nm. This result unequivocally demonstrates that this enzyme is not a flavoprotein and possesses none of the common electron carriers to mediate the electron transfer from NADH to O_2. The nature of the oxygen metabolite was determined to be H_2O_2. Since the ratio of NADH oxidized to H_2O_2 produced is approximately one, this enzyme-catalyzed NADH oxidation is clearly a two-electron redox process overall. A variety of alternate electron acceptors were examined to test their competence as oxidants for this enzymatic reaction. It was found that compounds that are well-known one electron oxidants could also serve as effective mediators. These observations strongly suggest that the enzyme-catalyzed H_2O_2 formation is not a direct two-electron reduction of molecular oxygen, but may instead be a one-electron reduction process followed by dismutation of the nascent superoxide ($2O_2^- + 2H^+ \longrightarrow H_2O_2 + O_2$).

The NADH-dependent O_2^--generating activity was later confirmed by measuring the superoxide dismutase inhibitable reduction of ferricytochrome c. Since the active enzyme is metal-free, the indisputable formation of O_2^- as the proximate reducing intermediate suggests the participation of an enzyme-bound organic cofactor mediating the obligatory $2e^-/1e^-$ conversion as electrons pass on from NADH to O_2 in the catalysis. Substantiating this proposition was the finding that the purified enzyme alone could accept two electrons from NADH stoichiometrically in the absence of any electron mediators. The observation of a characteristic free radical signal (g=2.002) in the EPR spectrum obtained anaerobically with a sample of the enzyme and NADH at 8 °K may also support the existence of an organic cofactor. Since this enzyme is expected to operate via a single mechanism despite its dual functions as a NADH oxidase and a 3,4-glucoseen reductase, the unique $2e^-/1e^-$ switching capability found for this enzyme provides, for the first time, compelling evidence that it may operate through a radical mechanism. Thus, the C-3 deoxygenation in the biosynthesis of ascarylose, and possibly the 3,6-dideoxyhexoses in general, may proceed with CO bond disruption followed by stepwise $1e^-/1e^-$ reduction.

"Cycloaddition Reactions of Allyl and Related Cations"
RO1 GM 35962 (Gassman, P.), University of Minnesota

Nature is extremely effective in generating a wide variety of monocyclic and polycyclic ring systems containing only carbon. The exact biosynthetic paths to many of these systems are unknown. Even in many cases where the biosynthetic paths are believed to be understood, scientists have been unable to duplicate the specificity of nature. During the past 2 years, Dr. Gassman and his colleagues have demonstrated that acyclic hydrocarbons can be converted to certain monocyclic and bicyclic hydrocarbons in high yield with high specificity at room temperature or below. These reactions accomplish two goals. First, they provide examples of cyclization reactions under mild conditions which will add to the synthetic repertoire of chemists and simplify the synthetic approaches to a variety of terpene type natural products. Second, they provide examples of intramolecular cyclizations, which had been ruled out as possible biosynthetic pathways, because of the high temperature and high pressures previously thought to be necessary in order to accomplish

such transformations. These results should prompt a rethinking of certain proposed biosynthetic routes.

Dr. Gassman and coworkers have found that protonation of certain acyclic tetraenes leads to the formation of reactive intermediates having a conjugated diene in one part of the molecule and an allyl cation in another portion of the molecule. The allyl cation can be viewed as a vinyl group bearing a powerful electron-withdrawing group (the carbocation). As a result, this vinyl group is transformed into an extremely powerful Diels-Alder dienophiles. This dienophile adds intramolecularly to the conjugated diene to form a bicyclic ring system. Many of these reactions are complete in seconds below room temperature. Thus, they serve as excellent models for nature. In this way, bicyclo[4.3.0]nonanes, bicyclo[4.4.0]decanes, and bicyclo[5.4.0]undecanes have been prepared in excellent yield and high specificity.

In an attempt to model successfully nature's formation of 10-, 11-, 12-, and 14-membered carbocyclic skeletons, Dr. Gassman and coworkers have shown that an alkoxylstabilized allyl cation can be trapped intramolecularly by a vinyl cyclopropane group to form 11-membered rings. In one case, a pair of diastereomers were formed in 92 percent yield at -23°C to ambient temperature over 17 hours, which illustrated the facile nature of this intramolecular cyclization.

"Synthetic Ionophores for Cation Regulation"
RO1 GM 36262 (Gokel, G.), University of Miami

Cation complexation is fundamental to many biological processes. In particular, the recognition and transport of alkali metal cations such as Na^+, K^+, and alkaline earth metals such as Ca^{++}, are especially important since these cations are ubiquitous. The problem of recognizing a cation is often conceptualized in size terms: if the guest (cation) is a specific size, it will fit (and therefore presumably be recognized) by the same size but complementary host. The fit is no doubt an important part of cation complexation and regulation, but the combination of how rapidly the cation is bound coupled with the strength of binding also plays a crucial role.

Dr. Gokel's cation regulation program is based on the notion that both kinetics and thermodynamics are crucial to the problem. The rates of cation binding and release define the equilibrium position of the reaction between a ligand and a cation. His group has designed a large variety of macrocyclic polyether compounds having one or more side arms. These compounds are generally flexible and therefore have fast cation complexation rates. They also enjoy reasonably fast release rates so that transport through a membrane is possible. This is important for cation transport since compounds having very slow release rates will be poor cation carriers even though they may have good selectivity and binding strength.

The lariat ether compounds were devised to be both selective and dynamic. They are based generally on the crown ether framework but have one or more flexible arms that can assist in solvating a ring-bound cation. As part of the present program, this team has prepared more than 200 novel structures containing one, two, or three side arms. Selectivity and binding strength have been studied in homogeneous solution and some of the factors affecting

selectivity are now better understood. They have been able, for example, to rank donors in their efficacy for binding Na^+, K^+, and Ca^{++}. They have also been able to dispel the longstanding myth that selectivity is controlled exclusively or even primarily by the "hole-size fit" notion. Cation binding strength, dynamics, and both enthalpic and entropic contributions to binding have been assessed for many structures. Solid state structures have been obtained for numerous compounds and this suggests the binding relations that exist in solution. They have developed a number of dynamic binders that have 20-fold selectivity for Ca^{++} over identically sized Na^+. They have prepared lariat ether compounds having steroid side arms which exhibit interesting and novel aggregation behavior and represent a new class of neutral liposomes. The work accomplished thus far lays the groundwork for a new thrust in the direction of synthetic cation-conducting channels.

"The Biological Chemistry of Sulfur and Selenium"
R01 GM 37000 (Rabenstein, D.), University of California, Riverside

Although the importance of sulfur, and more recently selenium, as essential elements is well established, many aspects of their biological chemistry remains uncharacterized at the molecular level. Dr. Rabenstein and his coworkers have made substantial progress in characterizing, at the molecular level, selected reactions of sulfur and selenium compounds in intact erythrocytes, plasma, and aqueous solution by nuclear magnetic resonance (NMR) spectroscopy. He and his colleagues have developed methods with which the water resonance and interfering hemoglobin resonances can be selectively eliminated from proton NMR spectra of intact erythrocytes. With the high sensitivity of their 500 MHz NMR spectrometer, they can obtain a spectrum from compounds present in erythrocytes at concentrations as low as 50 micromolar in as little as 15 seconds of instrument time. With this high sensitivity, they have been able to monitor changes in the redox state of intracellular thiol compounds after the erythrocytes are subjected to oxidative stress. Endogenous thiol compounds and thiol-containing drug molecules, e.g. penicillamine, which were introduced into the erythrocytes by incubation have been studied. With the high chemical shift dispersion and high sensitivity of their NMR spectrometer, they were able to detect well-resolved resonances for intracellular glutathione, oxidized glutathione, penicillamine, penicillamine disulfide, and penicillamine-glutathione mixed disulfide in penicillamine-containing erythrocytes which had been subjected to oxidative stress. They found that, after glucose is added to the stressed erythrocytes, the oxidized glutathione and the penicillamine-glutathione mixed disulfide are reduced by an enzyme-catalyzed process, whereas penicillamine disulfide is not. The results of this study demonstrate that proton NMR spectroscopy is a powerful method for studying the intracellular oxidation/reduction chemistry of small thiol-containing compounds.

Dr. Rabenstein and his coworkers have also discovered that thiol/disulfide interchange reactions, which are important pathways for the metabolism of thiol-containing compounds, can be quantitatively characterized by NMR spectroscopy. They have characterized reactions involving glutathione, cysteine, coenzyme A, captopril, and homocysteine. In a parallel study which has as a major objective the elucidation of the chemistry which occurs at the active site of the seleno-enzyme glutathione peroxidase, they have

quantitatively characterized a selenol/diselenide exchange reaction. They have found that selenol/diselenide exchange is some 10^7 times faster than thiol/disulfide exchange at physiological pH. This is an amazing difference in reactivity for these two otherwise very similar elements, which suggests that nature may have chosen selenium rather than sulfur for the active site of glutathione peroxidase because of the combined effects of selenolate being both a better nucleophile and a better leaving group in nucleophilic displacement reactions.

"Structural Models for Molybdoenzymes"
R01 GM 37773 (Enemark, J.), University of Arizona

Molybdoenzymes are essential in sulfur metabolism and in several other biological processes. However, the detailed structure of the molybdenum center of these enzymes is still not known. Dr. Enemark and his colleagues have been synthesizing model molybdenum compounds which incorporate chemical and structural features proposed to be present in the molybdenum centers of enzymes. Previous work by his group and others have shown that at least two sulfur atoms must be bound to the molybdenum in order mimic the chemical reactivity and spectroscopic properties of molybdenum in sulfite oxidase and related enzymes. Most recently they have synthesized a series of oxomolybdenum(V) compounds with the same donor atoms but varying chelate ligand backbones. It was found that the size of the chelate ring and of the alkyl substituents on the chelate ring backbone of $LMoO[S\text{-}(CH_2)_n\text{-}S]$ and $LMoO[O\text{-}(CH_2)_n\text{-}O]$ compounds can lead to changes in the reduction potential of more that 0.2 volts per CH_2 unit. The detailed structural changes at the molybdenum atom that accompany the changes in ring size are not yet known. The protecting ligand (L) constrains the molybdenum to six-coordinate fac-stereochemistry, but changes in the torsional angles of the coordinated heteroatom can result from changes in chelate ring size. Such torsion angle changes provide a promising mechanism for transmitting seemingly remote differences in the ligand backbone to the reduction potentials of the metal center.

Finally, the significant changes in molybdenum reduction potentials that are observed upon changing the ligand skeleton may be important for understanding the properties of the molybdenum cofactor of enzymes, which is thought to possess a five-membered chelate ring with two side chains.

"The Plane Facts on the Mode of Action of Bicyclomycin"
R01 GM 37934 (Kohn, H.), University of Houston

Bicyclomycin is a clinically useful antibiotic possessing a diverse spectrum of biological activity. This structurally unique drug exhibits moderate to potent activity against several Gram-negative bacteria including Escherichia coli, Klebsiella, Shigella, Salmonella, Citrobacter, Enterobacter cloacae, and Neisseria gonorrheae. Its emerging importance for the treatment of nonspecific diarrhea in humans and bacterial diarrhea in calves and pigs has led to the commercial introduction of this drug under the trade name: Bicozamycin. Unfortunately, efforts aimed at elucidating the mechanism of action of bicylomycin at the molecular level both in vitro and in vivo have

not paralleled the introduction and use of the drug. Little is known concerning the drug activation process, the energetics of key steps, and the interaction of the antibiotic with the receptor site. Both chemical and enzyme-mediated pathways have been invoked to account for the mode of action of the drug.

The major goals of studies instituted in Dr. Kohn's laboratory have been the elucidation of the chemical pathways for drug activation, and the determination of the minimal conditions needed for activation under conditions which approximate the biological process. Significantly, Drs. Kohn and Abuzar have discovered that drug activation proceeds over a wide pH range. Of particular interest, recent studies have demonstrated that treatment of bicyclomycin with alkyl mercaptans, cysteine derivatives, and select secondary amines in tetrahydrofuran-water mixtures at near neutral 'pH' led to the stereoselective formation of a single compound. The reactions proceeded at the exomethylene group in bicyclomycin with the loss of ammonia. These transformations represented the first documented examples of the activation of the drug under mild conditions. Furthermore, the observation that bicyclomycin binds to cysteine derivatives has commanded special attention in light of previous suggestions that sulfhydryl-containing proteins present within the peptidoglycan assembly of the growing bacteria are the primary targets for the antibiotic. Finally, the exomethylene functionalized product has been shown to undergo further binding with lysine derivatives to generate bis-alkylated adducts. These discoveries have permitted Dr. Kohn and coworkers to suggest that the antibiotic functions as a sequential alkylating agent and to propose a new pathway for its mode of action.

"Antibodies as Catalysts"
R29 GM 38273 (Hilvert, D.), Scripps Clinic and Research Foundation

The mammalian immune system is the most prolific source of specific receptor molecules known. Recently, scientists have demonstrated that this marvelously diverse system can be exploited to prepare antibody proteins with tailored catalytic activities. This approach to enzyme design involves synthesizing compounds that mimic the transition state of a particular reaction, eliciting immune responses against such substances, and characterizing the specific antibodies generated.

Dr. Hilvert is utilizing this new technology to develop antibody catalysts for concerted electrocyclic reactions, including Claisen rearrangements and Diels-Alder cyclizations. Since a presumably low probability exists for generating an effective constellation of multiple catalytic groups (e.g., general acids, general bases, nucleophiles) in the binding site of an antibody during immunization, reactions that are shape- rather than chemoselective should be especially good targets for catalysis. Concerted reactions generally do not require chemical catalysis but should be particularly sensitive to the principal catalytic effects antibodies are likely to impart, i.e. induced strain and proximity. Moreover, these processes are of enormous theoretical and practical interest, especially for enhancing our understanding of how enzymes work and for synthesizing biologically active molecules.

The conversion of (-)-chorismate into prephenate is an example of a biologically relevant 3,3-sigmatropic rearrangement. In plants and lower

organisms this transformation, catalyzed by the enzyme chorismate mutase, is the committed step in the biosynthesis of the aromatic amino acids phenylalanine and tyrosine. Dr. Hilvert used a transition state analog inhibitor for the natural enzyme to prepare antibodies with chorismate mutase activity. One of the antibodies catalyzed the rearrangement with a rate acceleration of more than two orders of magnitude compared to the uncatalyzed process. Saturation kinetics were observed, and at 25° C the values of k_{cat} and K_m were 1.2 X 10^{-3} s^{-1} and 51 μM, respectively. The transition state analog was shown to be a competitive inhibitor of the reaction with K_i equal to 0.6 μM. The antibody also exhibited high enantioselectivity, accepting only the (-)-isomer as a substrate.

These results represent the first example of an antibody-promoted carbon-carbon bond-forming reaction and demonstrate the feasibility of catalyzing concerted electrocyclic processes with antibodies. The strategy utilized in Dr. Hilvert's experiment is now being extended to the development of other antibodies that are shape-selective rather than chemoselective. Such molecules will be valuable as tools to study the importance of strain and proximity in protein-catalyzed reactions or for the synthesis of complex natural products. Since the efficacy of many drug molecules depends on the precise configuration of bonds around a single atom, the superb stereo-specificity of catalytic antibodies, similar to that of natural enzymes, should be of great practical value.

"Chemical Interactions of Coral Reef Invertebrates"
R29 GM 38624 (Paul, V.), University of Guam

Research in Dr. Paul's laboratory is directed towards understanding the chemical defenses of marine organisms on coral reefs as a way of investigating bioactivities of marine natural products. Dr. Paul and her coworkers have developed a field-oriented approach to identifying bioactive natural products from marine invertebrates on Guam. They use a field assay that incorporates extracts and isolated metabolites into a carrageenan-based diet to test the effects of these compounds as feeding deterrents toward natural populations of coral reef fishes. These methods allow for a bioassay-guided approach to isolating natural products from marine invertebrates that function as chemical defenses. The metabolites are further tested for their pharmacological activities in Dr. Bob Jacob's laboratory. Thus, they can compare feeding deterrent properties of natural products with potential biomedical activities.

Dr. Paul's research tries to understand ecological factors associated with invertebrates that produce biologically active natural products. For example, soft corals produce two types of defenses against predators, secondary metabolites and calcified sclerites. Concentrations of secondary metabolites are much greater in the top portions of the colonies, while sclerite concentrations are much higher in the bases. Top portions are more accessible to predators but at the same time, high concentrations of secondary metabolites make the tops better chemically defended. Similarly, high concentrations of secondary metabolites are found in soft corals and gorgonian corals that have low sclerite concentrations or very small, powdery sclerites. The soft coral Sinularia maxima on Guam has relatively low concentrations of very fine sclerites but produces a potent cembranoid diterpenoid that functions as a chemical defense against predatory fishes. This same

metabolite shows potent anti-inflammatory activity in pharmacological assays that test for inhibition of phorbol-induced inflammation of the mouse ear. This metabolite is one of the best natural feeding deterrents and anti-inflammatory agents that has been isolated to date.

Other potent feeding deterrents that they have isolated from sponges and ascidians (tunicates) have also been very active in other pharmacological assays. Patellazole B from the tunicate Lissoclinum Patella and laulimalide from the sponge Hyattella sp. were both potent cytotoxins against the KB cell line, and extracts of each organism were effective feeding deterrents in field assays. Many other extracts and isolated metabolites have been investigated in these bioassays and the relationships between natural chemical defenses and cytotoxicity and anti-inflammatory activities are good. These methods may provide a useful way of relating chemical ecology with biomedical activities of marine natural products.

"Novel Ring-Enlargement Chain Extension Reaction"
RO1 GM 39825 (Dowd, P.), University of Pittsburgh

In the course of their investigations into the mechanism of action of vitamin B_{12}, Dr. Paul Dowd and his colleagues discovered an unusual rearrangement reaction involving migration of an α-keto ester group to a methylene radical center. The essentials of a novel ring expansion/chain extension reaction were contained in these experiments. The first ring expansion experiments took the form of the alkylation of five-, six-, and seven-membered cyclic β-keto esters, readily available from the Dieckmann condensation, with methylene dibromide or diiodide. Ring expansion to the respective six-, seven-, and eight-membered β-keto esters occurred smoothly in high yields upon treatment with tri-γ-butyltin hydride in refluxing benzene. This set the stage for development of a new synthetic strategy for ring expansion.

The next step involved extension of the reaction to include ring expansion by larger increments. In this fashion, attachment of three- and four-carbon halogen-bearing sidechains to the keto ester followed by treatment with tri-n-butyltin hydride resulted in ring expansion by three and four carbon atoms. Thus, entry into the medium-sized rings of 9, 10, and 11 carbons becomes straightforward. The method is compatible with ketone and ester groups, so it encourages inclusion of useful functionality for further synthetic elaboration.

Current efforts involve extension of the method to heterocyclic compounds, where there are special problems to overcome. For example, a new precursor was required because the bromomethyl sidechains undergo undesirable side reactions. Accordingly, the selenophenylmethyl group was used as the precursor to the free radical. Using this approach N-benzylpiperidone carboxylate undergoes smooth ring expansion to the corresponding seven-membered azepinone. Likewise, ring expansions of tetrahydrofuryl and tetrahydrothiophyl keto esters were readily affected.

The ring expansion method has also been applied to large rings. Dr. Dowd's method allows expansion of 12- to 13-, 14- to 15-, and 15- to 16-membered rings. An attractive target molecule in this series is the 15-membered,

naturally occurring perfume constituent muscone. By attaching a 2-methyl-1-iodopropyl side chain to readily available cyclododecalnone and treating with tri-n-butyltin hydride, (±)-muscone was produced in 20 percent yield. The optically active sidechain precursor is commercially available, so enantiomerically pure muscone is also accessible by this route.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1990— 717 - 022/06089

NIH Library, Building 10
National Institutes of Health
Bethesda, Md. 20892

NIH LIBRARY
3 1496 00371 2521

Administrative Document
Morch 1990